LES MATHS EN PHYSIQUE

La physique à travers le filtre des mathématiques

avec éléments d'analyse numérique

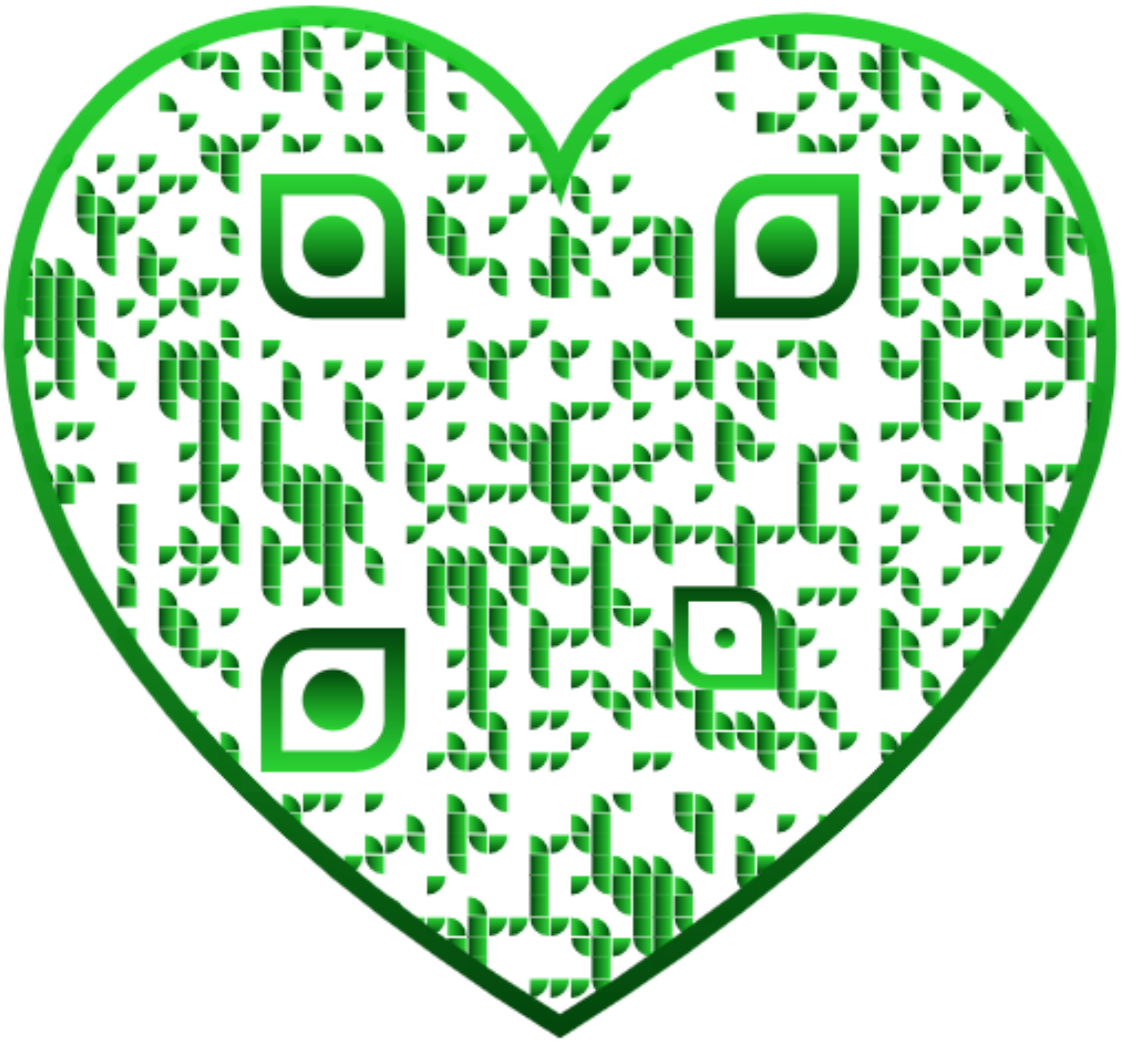
Cours et applications

Jean-Pierre Provost
Professeur à l'université de Nice-Sophia Antipolis

Gérard vallée
Maître de conférences à l'université de Nice-Sophia Antipolis

3e édition

DUNOD



Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.
Le Code de la propriété intellectuelle du 1er juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.
Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).

© Dunod, Paris, 2011
ISBN 978-2-1005-5933-6
© Dunod, Paris, 2006 pour la 2e édition

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).
Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

Liste des abréviations **vi**

Table des sujets de physique **vii**

Avant-propos **xi**

1 Nombres réels ; grandeurs physiques ; dimensions **1**
1.1 Grandeurs physiques ; continuité ; paramétrages additifs 1
1.1.1 Survol « physique » des ensembles $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ 1
1.1.2 Paramétrage additif des lois de composition ; logarithmes 4
1.1.3 Fonction et notation exponentielles ; applications 6
1.1.4 Mesure additive du désordre microscopique ; grands nombres et entropie ; exemples ; irréversibilité 8
1.2 Caractère algébrique des grandeurs physiques 10
1.2.1 Pensée « naïve » et pensée algébrique 10
1.2.2 Conventions et lois de l'électricité 12
1.3 Grandeurs physiques et dimensions 15
1.3.1 Changements d'unités et invariance des lois physiques 15
1.3.2 Applications et limites de l'analyse dimensionnelle 18

2 Nombres et notation complexes ; plan euclidien **22**
2.1 Calculs avec les nombres complexes 22
2.1.1 Règles de calcul ; exponentielle imaginaire ; fonctions complexes . . 22
2.1.2 « Théorème fondamental de l'algèbre » et applications 25
2.2 Plan complexe et transformations associées 26
2.2.1 Plan complexe et plan cartésien ; produit scalaire ; aire 26
2.2.2 Transformations dans le plan complexe 28
2.3 Etude de courbes et de mouvements plans 31
2.3.1 Mouvements et courbes en coordonnées polaires 31
2.3.2 Coniques en coordonnées polaires et cartésiennes ; foyers 32
2.3.3 Mouvement de Kepler 36
2.3.4 Mouvement harmonique ; vecteurs tournants 38
2.4 Notation complexe en physique classique 39
2.4.1 Signaux réels et complexes 39
2.4.2 Systèmes entrée-sortie ; fonctions de transfert et impédances 40
2.4.3 Signaux modulés ou quasi-monochromatiques 42
2.5 Applications à l'optique ondulatoire 42
2.5.1 Interférences 42
2.5.2 Diffraction en lumière monochromatique 45
2.5.3 Polarisations 46

3 Espace ; symétries ; calcul vectoriel **51**
3.1 Symétrie, invariance et relativité 52
3.1.1 Groupes de symétrie et invariance 52
3.1.2 Le groupe de symétrie de la physique 53
3.1.3 Symétries spatiales (présentation « expérimentale ») 55
3.1.4 Transformation des grandeurs et des champs physiques 59
3.2 Calcul vectoriel ; applications 60
3.2.1 Produit scalaire, produit vectoriel et produit mixte ; pseudovecteurs 60
3.2.2 Equations de plans ; fréquences spatiales ; réseaux 64
3.2.3 Différentielles de chemins ; effet Doppler ; lois de Descartes 66
3.2.4 Vecteurs surface ; flux de grandeurs 69
3.2.5 Sphère, angle solide et applications 71
3.2.6 Géométrie sphérique ; notion de transport parallèle 74
3.3 Vecteurs tournants ; mécanique du solide 76
3.3.1 Vecteurs tournants ; changements de référentiels 76
3.3.2 Référentiel du centre de masse 78
3.3.3 Cinématique et dynamique d'un corps solide 80
3.4 Systèmes physiques possédant des symétries 83
3.4.1 Schéma général ; principe de Curie 83
3.4.2 Symétries de translation ; lois de Descartes et généralisations . . . 84
3.4.3 Symétries de rotation et symétries discrètes ; applications en électromagnétisme et en acoustique 85

4 Calcul et physique linéaires ; relativité ; quantique **87**
4.1 Espaces vectoriels 87
4.1.1 Définitions ; changements de bases ; applications linéaires 87
4.1.2 Structures métriques ; fonctions orthonormées 90
4.1.3 Formes quadratiques et antisymétriques ; volume 91
4.2 Calcul matriciel 94
4.2.1 Bases du calcul matriciel ; lien avec le calcul vectoriel 94
4.2.2 Matrices $n \times n$; exponentielle ; trace ; déterminant ; inverse 95
4.2.3 Spectre d'une matrice $n \times n$; vecteurs propres ; exemples 98
4.2.4 Matrices de Pauli ; groupe de symétrie de la physique ; spineurs . . 101
4.2.5 Groupe de rotation et classification des grandeurs physiques 103
4.3 Applications en physique classique 105
4.3.1 Déformations et contraintes ; élasticité ; viscosité 105
4.3.2 Optique matricielle des systèmes centrés 108
4.3.3 Relativité d'Einstein et quadrivecteurs 111
4.4 Physique quantique et espaces vectoriels 117
4.4.1 Cadre général ; états quantiques ; moyennes ; évolution 118
4.4.2 Dynamique des systèmes à deux états ; transitions quantiques . . . 120
4.4.3 Fonctions d'ondes ; E.D.P. de Schrödinger ; états gaussiens 124
4.4.4 Oscillateur harmonique ; champ électromagnétique 127
4.4.5 E.D.P. relativistes de Klein-Gordon, Weyl, Dirac ; champs quantiques ; particules et antiparticules 130

5 Fonctions d'une variable ; analyse des signaux 133
5.1 Savoir-faire concernant les fonctions 134
5.1.1 Graphe et informations sur une fonction 134
5.1.2 Dérivation, développements limités : principaux résultats 137
5.1.3 Intégration ; cas des fonctions piquées ou rapidement oscillantes . . 139
5.1.4 Concavité de l'entropie ; travail maximum ; transitions de phase . . 141
5.2 Opérations sur les fonctions ; analyse de Dirac 145
5.2.1 Opérations sur les fonctions 145
5.2.2 Impulsion de Dirac (« fonction delta ») ; exemples mécaniques . . . 147
5.2.3 Analyse de Dirac ; réponse impulsionnelle ; convolution ; filtrage . . 150
5.3 Transformation de Fourier ; analyse de Fourier 151
5.3.1 Décomposition de Fourier ; spectre d'un signal ; formule de Poisson 152
5.3.2 Propriétés : dualité temps fréquence 154
5.3.3 Transformée de Laplace 157
5.3.4 Signaux stationnaires ; signaux chaotiques ; langage probabiliste . . 158
5.4 Optique de Fourier ; filtrage optique 161
5.4.1 Décomposition en ondes planes ; filtrage 161
5.4.2 Illustrations optiques de la transformée de Fourier 163

6 Equations différentielles ; systèmes dynamiques 166
6.1 Systèmes dynamiques et espace de phase 167
6.1.1 Définitions ; propriétés générales 167
6.1.2 Exemples de systèmes dynamiques et de leurs portraits de phase . 170
6.2 Equations linéaires stationnaires ; modes propres ; stabilité 174
6.2.1 Equations du premier et du second ordre sans et avec second membre174
6.2.2 Cas général ; modes propres ; oscillateurs couplés 177
6.2.3 Stabilité et instabilité d'un système dynamique linéaire stationnaire 180
6.3 Dix équations vectorielles classiques de la physique 182
6.4 Equations différentielles linéaires à coefficients variables 187
6.4.1 Quatre exemples ; matrices de transfert 187
6.4.2 Ondes stationnaires ; états liés ; quantification 190
6.4.3 Ondes propagatives ; réflexion, transmission, adaptation d'impédance191
6.4.4 Equations avec paramètres périodiques ; théorème de Floquet-Bloch 193
6.4.5 Equations d'amplitude ; approximation adiabatique 195
6.5 Oscillateurs non linéaires 197
6.5.1 Oscillateurs linéairement stables faiblement non linéaires 197
6.5.2 Oscillateurs linéairement instables ; exemple de Van der Pol ; bifurcations de Hopf et d'un cycle limite 200

7 Fonctions de plusieurs variables ; analyse vectorielle 204
7.1 Calcul différentiel 204
7.1.1 Développement de Taylor ; différentielles ; variation seconde ; extremum ; E.D.P. simples 204
7.1.2 Dérivées spatiales de champs scalaires et vectoriels 208
7.1.3 Dérivées temporelles et applications hydrodynamiques 212
7.1.4 Changements de variables ; règles de calcul 214
7.2 Calcul intégral 216

7.2.1 Intégration à n dimensions ; jacobien ; cas des très grandes dimensions 216
7.2.2 Formes différentielles ; formules de Stokes et Ostrogradski 219
7.2.3 Analyse vectorielle ; frontières et champs dépendant du temps ; loi de Lenz ; hydrodynamique 224
7.2.4 Bilans de grandeurs ; applications aux milieux continus 227
7.3 Applications à la mécanique et à l'optique géométrique 231
7.3.1 Mécanique et fonctions énergie potentielle 231
7.3.2 Optique géométrique : rayons ; surfaces d'onde ; caustiques et problèmes d'extremum 234
7.4 Applications à la thermodynamique 238
7.4.1 Rôle clé de l'entropie ; équations d'état ; coexistence de phases . . . 239
7.4.2 Potentiels thermodynamiques et équilibres 242
7.5 Applications à l'électromagnétisme 245
7.5.1 Formulation intégrale ; champs statiques ; milieux 245
7.5.2 Potentiel scalaire et potentiel vecteur ; bilans d'énergie et de quantité de mouvement ; A.R.Q.S 248
7.5.3 Calculs avec des densités microscopiques ; rayonnement 250

8 Equations aux dérivées partielles ; propagation ; diffusion 253
8.1 Chaines de systèmes dynamiques couplés ; limite continue 253
8.1.1 Chaines d'oscillateurs ; rôle des conditions aux limites ; phonons . . 254
8.1.2 Limite continue ; cordes vibrantes ; lignes électriques ; hydrodynamique ; impédances 256
8.2 Solutions de quelques E.D.P. dynamiques 261
8.2.1 E.D.P. linéaires à coefficients constants et solutions ondes planes . 261
8.2.2 Equations de diffusion et de propagation ; fonctions de Green ; ondes stationnaires 267
8.2.3 Six E.D.P. liées à une loi de conservation 272
8.2.4 Trois exemples d'E.D.P. non linéaires ; ondes solitaires 273
8.3 E.D.P. « spatiales » impliquant l'opérateur laplacien 274
8.3.1 Exemples et analogies physiques ; conditions aux limites 274
8.3.2 Unicité des solutions ; identité de Green ; séparation des variables . 276
8.3.3 Equation de Laplace $\Delta f = 0$ dans le plan et fonctions d'une variable complexe ; applications hydrodynamiques 279

9 Principes variationnels 283
9.1 Exemples historiques ; formalismes de Lagrange et de Hamilton 283
9.1.1 Principes de Fermat, Maupertuis, Lagrange 283
9.1.2 Principe de Hamilton dans l'espace de phase 286
9.1.3 Equations d'Euler-Lagrange ; symétries et lois de conservation ; E.D.P. d'Hamilton-Jacobi 287
9.1.4 Equations de Hamilton et géométrie de l'espace de phase 289
9.2 Principes de moindre action et généralisation des mouvements inertiels . . 290
9.2.1 Collisions et introduction de la masse inertielle 290
9.2.2 Particules chargées et interactions électromagnétiques 291
9.2.3 Temps propre et gravitation ; courbure de l'espace-temps 294
9.3 Champs et principes de moindre action 297

9.3.1 Equations de Maxwell ; action de Schwarzschild 297
9.3.2 Equation de Schrödinger ; théorème adiabatique 298
9.3.3 Equations d'Einstein de la gravitation ; action de Hilbert 299

10 Probabilités ; processus aléatoires 300
10.1 Langage des probabilités 300
10.1.1 Grandeurs aléatoires et raisonnements logiques ; conditionnement . 301
10.1.2 Probabilités ; lois de probabilité 302
10.1.3 Grandeurs moyennes ; moments ; corrélations 307
10.2 Origine et discussion de quelques lois importantes en physique 308
10.2.1 Théorème de la limite centrale et lois gaussiennes 308
10.2.2 Loi binomiale et loi de Poisson 310
10.2.3 Loi de Boltzmann ; généralisations ; statistiques quantiques ; réponse linéaire et fluctuations 312
10.2.4 Estimation et lois de χ^2 (khi-deux) 315
10.3 Processus aléatoires 317
10.3.1 Marche aléatoire ; processus de diffusion ; mouvement brownien . . 317
10.3.2 Processus de Markov ; probabilités de transition ; bilan détaillé . . 319
10.3.3 Processus stationnaires ; théorème de Wiener-Khintchine ; ergodicité 321

11 Analyse numérique ; physique discrète 323
11.1 Discrétisation 324
11.1.1 Représentation des nombres ; erreurs ; stabilité numérique 324
11.1.2 Dérivation et intégration ; extrapolation de Richardson 326
11.1.3 Nombres aléatoires ; méthode de Monte-Carlo 327
11.2 Résolution numérique d'E.D. et d'E.D.P. 328
11.2.1 Systèmes dynamiques ; schémas d'Euler et de Runge Kutta 328
11.2.2 E.D. avec conditions aux limites ; méthode de tir 330
11.2.3 E.D.P. avec conditions initiales : propagation, diffusion 331
11.3 Approximation de fonctions ; interpolation ; moindres carrés ; méthode de Bézier 334
11.3.1 Approximations polynomiales (Tchebytchev, B-splines...) 334
11.3.2 Interpolation de Lagrange et par "cubic-splines" 337
11.3.3 Méthode des moindres carrés 339
11.3.4 Méthode de Bézier 341
11.4 Résolution d'équations, d'E.D. et d'E.D.P. linéaires 343
11.4.1 Equations linéaires régulières et singulières 343
11.4.2 E.D. et E.D.P. linéaires ; méthodes spectrales ; éléments finis 347
11.5 Recherche de minima 350
11.5.1 Méthodes utilisant le gradient 350
11.5.2 Méthodes du simplex et du recuit simulé 352

Index 354

Liste des abréviations

A.R.Q.S. approximation des régimes quasi-stationnaires.

c.d.m. centre de masse.

C.I. condition initiale.

C.L. condition aux limites.

E.D. équation différentielle.

- E.D.L. équation différentielle linéaire.
- E.D.L.S. équation différentielle linéaire stationnaire.

E.D.P. équation aux dérivées partielles.

E.V. espace vectoriel.

L.C. loi de conservation.

S.D. système dynamique.

- S.D.L. système dynamique linéaire.
- S.D.L.S. système dynamique linéaire stationnaire.

T.F. transformée de Fourier.

v.a. variable aléatoire.

Table des sujets de physique

1. Généralités

1.1 *Caractère algébrique des grandeurs physiques* : Chapitre 1, pages 10-15.

1.2 *Dimensions ; analyse dimensionnelle* : Chapitre 1, pages 15-21.

1.3 *Grandeurs (pseudo)scalaires, (pseudo)vectorielles, quadrupolaires* : Chapitre 3, pages 59-63 et Chapitre 4, pages 103-104.

1.4 *Notation complexe* : Chapitre 2, pages 39-42.

1.5 *Statistiques de Gauss et de Poisson* : Chapitre 10, pages 308-312.

1.6 *Loi du χ^2, estimation* : Chapitre 10, pages 315-316.

2. Mécanique classique

2.1 *Etude de mouvements*
- Coordonnées polaires : Chapitre 2, pages 31-33.
- Mouvements de Kepler, harmonique, uniformément accéléré, de précession, d'une particule chargée dans un champ magnétique, pendule de Foucault : Chapitre 2, pages 36-38 et Chapitre 6, pages 182-187.
- Portraits de phase de mouvements à une dimension : Chapitre 6, pages 169-174.

2.2 *Oscillateurs*
- Oscillateurs linéaires amortis : Chapitre 6, pages 174-175.
- Oscillateurs paramétriques, approximation adiabatique : Chapitre 6, pages 173-174 et pages 193-196.
- Oscillateurs non linéaires : Chapitre 6, pages 171-174 et pages 197-203.
- Oscillateurs linéaires couplés : Chapitre 4, pages 92 et 100 et Chapitre 6, pages 178-180.
- Chaîne d'oscillateurs, limite continue, corde vibrante : Chapitre 8, pages 253-260.

2.3 *Mécanique et principes variationnels*
- Principes de Lagrange et de Hamilton : Chapitre 9, pages 283-287.
- Equations d'Euler-Lagrange, symétries et lois de conservation : Chapitre 9, pages 287-289.
- Introduction de la masse inertielle et des interactions électromagnétiques : Chapitre 9, pages 290-293.

2.4 *Autres sujets*
- Changements de référentiels, référentiel du centre de masse, problème des deux corps : Chapitre 3, pages 77-80.
- Cinématique et dynamique du solide : Chapitre 3, pages 80-82.
- Energie potentielle d'un système de points matériels et en électromagnétisme : Chapitre 7, pages 231-234.
- Résonance magnétique nucléaire : Chapitre 6, pages 177-178.

3. Hydrodynamique et milieux continus

3.1 *Déformations et contraintes ; élasticité, viscosité* : Chapitre 4, pages 105-108.

3.2 *Dérivée en suivant le mouvement ; théorèmes de l'hydrodynamique* : Chapitre 7, pages 212-214 et 226-227.

3.3 *Flux de masse, de quantité de mouvement, d'énergie ; bilans* : Chapitre 3, pages 70-71 et Chapitre 7, pages 227-231.

3.4 *Autres sujets*
- Equations adimensionnées : Chapitre 1, page 20.
- Ecoulements irrotationnels incompressibles : Chapitre 8, pages 274-277 et pages 279-282.

4. Thermodynamique et physique statistique

4.1 *Entropie*
- Introduction statistique : Chapitre 1, pages 8-10.
- Propriétés des fonctions $S(U)$ et applications (thermalisation, travail maximum, transitions solide-liquide et paramagnétique-ferromagnétique) : Chapitre 5, pages 141-144.
- Propriétés des fonctions $S(U, V)$ et applications (gaz parfait, fluide de Van der Waals, corps noir) : Chapitre 7, pages 238-242.

4.2 *Potentiels thermodynamiques ; équilibres, déplacements d'équilibres, paramètre d'ordre* : Chapitre 7, pages 242-243.

4.3 *Loi de Boltzmann ; réponse linéaire, fluctuations ; statistiques quantiques* : Chapitre 10, pages 312-315.

4.4 *Phénomènes de diffusion et de transport*
- Marche au hasard, limite continue : Chapitre 10, pages 317-319.
- Equations de diffusion, de Fokker-Planck, de Boltzmann : Chapitre 8, page 262, pages 267-268 et pages 271-272.

5. Electricité et électromagnétisme

5.1 *Circuits électriques*
- Conventions et lois de l'électricité : Chapitre 1, pages 12-15.
- Forces exercées sur les circuits : Chapitre 7, pages 231-234.
- Limite continue, lignes électriques : Chapitre 8, pages 257-259.

5.2 *Equations de Maxwell.* Formulation intégrale, approximation des régimes quasi-stationnaires, milieux diélectriques et magnétiques, équations de propagation, rayonnement : Chapitre 7, pages 245-252. Formulation variationnelle : chapitre 9, page 297.

5.3 *Autres sujets*
- Utilisation des symétries : Chapitre 3, pages 83-86.
- Conducteurs en mouvement, loi de Lenz : Chapitre 7, page 226.
- Images électriques : Chapitre 8, pages 276-277.
- Milieux diélectriques anisotropes : Chapitre 8, pages 262-263.
- Invariance relativiste : Chapitre 4, page 117.

6. Relativité

6.1 *Groupe de symétrie et invariance des lois physiques*
- Généralités : Chapitre 3, pages 51-53.
- Transformations de Lorentz, approximation galiléenne : Chapitre, 3 pages 53-55.

6.2 *Relativité d'Einstein*
- Temps propre, quadrivecteurs : Chapitre 4, pages 111-115.
- Mécanique-électromagnétisme : Chapitre 4, pages 115-117 et chapitre 9 pages 290-293 et pages 295-296.

6.3 *Relativité et gravitation*
- Géodésiques, courbure, métrique de Schwarzchild : Chapitre 9, pages 294-297.
- Equations d'Einstein : Chapitre 9, page 299.

7. Optique

7.1 *Optique géométrique*
- Principe de Fermat, lois de Descartes, théorème de Malus : Chapitre 3, pages 67-68 et Chapitre 9, pages 283-284.
- Optique matricielle des systèmes centrés : Chapitre 4, pages 108-111 et Chapitre 6, pages 187-190.
- Etendue optique et luminance : Chapitre 3, pages 73-74.
- Caustiques et aberrations : Chapitre 7, pages 234-238.

7.2 *Interférences*
- Interférences en lumière monochromatique : Chapitre 2, pages 42-45 et Chapitre 3, pages 64 et 67.
- Localisation : Chapitre 3, page 68.
- Interférences en lumière non monochromatique : Chapitre 2, pages 44-45 et Chapitre 5, page 159.

7.3 *Diffraction de Fresnel et de Fraunhofer*
- Principe d'Huygens-Fresnel : Chapitre 2, pages 45-46 et Chapitre 5, page 150.
- Filtrage optique ; optique de Fourier : Chapitre 5, pages 161-165.

7.4 *Lumière polarisée ; états de polarisation, interférences* : Chapitre 2, pages 46-50.

8. Ondes

8.1 *Ondes planes*
- Fréquences spatiales ; lois de Descartes et des réseaux, formule de Bragg : Chapitre 3, pages 64-65 et pages 84-85.
- Réflexion, transmission, adaptation d'impédance : Chapitre 6, pages 191-192 et Chapitre 8, page 260.
- Relations de dispersion ; paquets d'ondes ; ondes guidées : Chapitre 8, pages 262-267.

8.2 *Approximation de l'optique géométrique ; surfaces d'onde et rayons* : Chapitre 7, pages 234-236.

8.3 *Equations de propagation à une et trois dimensions*
- Solution générale ; fonctions de Green ; ondes stationnaires : Chapitre 8, pages 267-272.
- Approximation de Born de la diffusion : Chapitre 8, page 278.

8.4 *Ondes non linéaires ; ondes de choc ; ondes solitaires* : Chapitre 7, pages 207 et 228 et Chapitre 8, pages 273-274.

9. Physique quantique

9.1 *Etats quantiques, amplitudes de probabilité, évolution* : Chapitre 4, pages 118-120.

9.2 *Dynamique des systèmes à deux états ; transitions quantiques* : Chapitre 4, pages 120-124.

9.3 *Equation de Schrödinger à une dimension*
- Fonctions d'onde ; oscillateur harmonique : Chapitre 4, pages 124-129.
- Puits et barrières de potentiel, potentiels périodiques : Chapitre 6, pages 190-194.

9.4 *Physique quantique relativiste*
- Matrices de Pauli et espace-temps ; spineurs : Chapitre 4, pages 101-103.
- Champs quantiques ; particules et antiparticules : Chapitre 4, pages 130-132.

9.5 *Autres sujets*
- Lois de conservation : Chapitre 3, page 85.
- Spineurs et moments cinétiques : Chapitre 4, pages 103-105.
- Emissions spontanée et induite ; coefficients d'Einstein : Chapitre 4, pages 129-130 et Chapitre 10, page 321.
- Statistiques de Fermi-Dirac et de Bose-Einstein : Chapitre 10, page 314.
- Théorème adiabatique ; distinction « travail-chaleur » ; phase de Berry ; transport quantique : Chapitre 9, pages 298-299.
- Diamagnétisme et paramagnétisme ; effet Aharonov-Bohm : Chapitre 9, page 293.

Avant-propos

Cet avant-propos est destiné à expliciter l'esprit qui a prévalu à la rédaction de cet ouvrage et à permettre au lecteur d'en tirer le meilleur profit.

Esprit et originalité de l'ouvrage

Ce livre a son origine dans l'élaboration de documents de travail pour "agrégatifs" de physique constitués de questions-exercices de mathématiques ayant un rapport direct avec le programme du concours. Sous sa forme actuelle, qui s'adresse à un public plus large, cet ouvrage reprend les principaux concepts et techniques mathématiques intervenant dans la physique enseignée en premier et second cycles universitaires, et les illustre par de nombreux exemples issus de toutes les branches de la physique. Son objectif est de mettre en avant l'intérêt des mathématiques pour la physique en montrant comment un même concept ou résultat mathématique intervient dans des domaines de physique parfois très éloignés les uns des autres, ce qui souligne l'essence mathématique de beaucoup d'idées physiques. Il ne s'agit donc ni d'un livre de "méthodes mathématiques pour la physique", avec définitions, lemmes et théorèmes, offrant un cadre rigoureux à un nombre limité d'applications physiques, ni d'un livre d'entraînement à la résolution de problèmes réduisant les mathématiques au rôle d'outils de calcul.

Quelques exemples

Des savoir-faire "typiques" de physicien, comme par exemple
1) l'établissement du signe correct dans une relation en raisonnant sur un cas particulier,
2) l'utilisation d'arguments dimensionnels pour rétablir des paramètres posés égaux à 1 lors du calcul d'une expression,
3) l'emploi d'arguments de symétrie pour déterminer la direction d'un champ vectoriel en un point,
reposent tous sur la propriété des lois physiques d'être invariantes par rapport à un groupe de transformations. S'il veut transmettre ce savoir le physicien se doit d'expliciter cette notion d'invariance, ce qui implique de préciser les transformations considérées et la manière dont elles agissent sur les grandeurs physiques. Ces informations sont non seulement très fondamentales mais aussi très pratiques.
Dans le premier cas ce sont le caractère algébrique des grandeurs physiques et l'invariance des lois par rapport aux changements de conventions de signe qui sont en jeu. L'exemple, donné au chapitre 1, d'un système électromécanique simple, avec 64 combinaisons de signe possibles, montre la relativité des conventions et la possibilité d'en choisir une sans perte de généralité (de même qu'on traite un problème dans un référentiel). Dans le second cas c'est l'invariance des lois par rapport aux changements d'unités qui est concernée. Cette propriété permet d'adimensionner des équations (ce qui simplifie beaucoup de calculs) et de faire des prédictions dimensionnelles (par exemple estimer des ordres de grandeurs). Quant aux arguments de symétrie, on montre par exemple au chapitre 3 que, conjuguée à la linéarité, l'invariance par rapport aux transformations les plus simples (les translations) conduit à la fois aux lois de Descartes (translations "continues"), aux formules des réseaux optiques et de Bragg pour les cristaux (translations "discrètes") et

aussi à la loi de conservation de la quantité de mouvement et de l'énergie (translations spatiales et temporelles) en physique quantique.

Beaucoup d'autres exemples relatifs aux liens mathématiques-physique pourraient être donnés. Ainsi au chapitre 5 on verra que ce qui se cache derrière une analyse de Fourier des fonctions est la notion physique de filtrage, derrière une analyse "à la Dirac" celle de réponse impulsionnelle et derrière une analyse "probabiliste" celle de corrélation.

Parties originales de l'ouvrage

En plus du caractère nouveau de son approche ce livre contient certaines présentations, discussions et démonstations originales (ou partiellement originales) comme :

- les introductions des dimensions (section 1.3.1), des groupes de symétrie (section 3.1.1), des bilans de grandeurs et ondes de choc à partir de l'exemple de la circulation automobile (section 7.2.4), de la masse d'inertie (section 9.2.1)...

- les discussions de la concavité de l'entropie dans des cas simples (section 5.1.4) et pour le gaz de Van der Waals (section 7.4.1), de la distinction entre signaux stationnaires chaotiques et quasipériodiques (section 5.3.3), du principe d'Huygens Fresnel en liaison avec la notion de filtrage (section 5.4.1), des volumes en grande dimension (section 7.2.1), des formules de Stokes et d'Ostrogradski pour des champs et des domaines dépendant du temps (section 7.2.3), de l'identité des vitesses de groupe et de l'énergie (section 8.2.1)...

- les démonstrations des lois de Descartes et de ses généralisations (section 3.4.2), de l'invariance relativiste de l'électromagnétisme (section 4.3.3), de la règle d'or de Fermi et de la loi exponentielle de désexcitation en quantique (section 4.4.2), du théorème adiabatique classique (section 6.4.5) et quantique (section 9.4.2), de la solution causale de l'E.D.P. de propagation à trois dimensions (section 7.2.2), de la formule de Biot et Savart sans les potentiels vecteurs (section 7.5.1), des lois de conservation à partir des principes variationnels (section 9.1.3)...

Organisation de l'ouvrage

L'ouvrage contient dix chapitres consacrés chacun à un grand domaine des mathématiques (nombres réels, complexes, espace et symétries, calcul linéaire, fonctions d'une variable, de plusieurs variables, équations différentielles, aux dérivées partielles, principes variationnels, probabilités) ainsi qu'aux applications physiques correspondantes.

Chaque chapitre commence par une brève introduction du sujet (souvent historique) permettant de motiver son étude ainsi que la façon de l'aborder. Les aspects les plus mathématiques (définitions, principaux résultats...) sont ensuite présentés dans les premières sections. Ces résultats font en général l'objet d'une démonstration rapide, certains calculs étant laissés au lecteur ; sinon l'idée de la démonstration est donnée. Ils ne sont donc jamais "parachutés" et leur appropriation par le lecteur est facilitée par la présence de nombreux exemples et éventuellement contre exemples.

Bien que l'imbrication mathématiques-physique soit permanente, les dernières sections d'un chapitre sont généralement consacrées à des domaines particuliers de physique utilisant les techniques et concepts introduits auparavant. C'est le cas par exemple de l'optique de Fourier (en liaison avec la transformation de Fourier), de la formulation intégrale de l'électromagnétisme (en liaison avec l'analyse vectorielle), de l'étude des déformations (en liaison avec le calcul matriciel)...

Au total le livre comporte plus d'une centaine de sous-sections traitant d'un thème précis, à dominante mathématique ou physique. Leur rédaction a été conçue de sorte à rendre chacune la plus autonome possible afin qu'un lecteur disposant d'un minimum de connaissances puisse s'y reporter directement (moyennant éventuellement des renvois, signalés dans le texte, à d'autres sous-sections).

A la table des matières usuelle a été ajoutée une "**table des sujets de physique**" destinée aux lecteurs qui souhaitent retrouver un sujet de physique connu dans le "découpage classique" par spécialités : Généralités, Mécanique classique, Hydrodynamique et milieux continus, Thermodynamique et physique statistique, Electricité et électromagnétisme, Relativité, Optique, Ondes, Physique quantique.

Différents niveaux de lecture

Le profit tiré de la lecture d'un livre dépend bien sûr de la culture du lecteur. Aussi les auteurs sont-ils conscients qu'ils ne peuvent s'adresser à chacun avec la même efficacité. Si il est vrai que les parties mathématiques sont en général facilement abordables car elles ne font appel qu'à des connaissances qui ne dépassent que rarement le niveau licence, les parties physiques demandent par contre de la part du lecteur une "participation active". En effet, l'esprit même du livre l'oblige à passer rapidement d'un domaine de physique à un autre ; cet exercice auquel les étudiants ne sont pas habitués est pourtant à la base de l'acquisition d'une véritable culture en physique.

Le lecteur doit être averti qu'il y a cependant différents niveaux de lecture de chaque chapitre ou même de chaque section. Si certaines parties ne seront pleinement profitables qu'à un étudiant de master ou même un enseignant possédant déjà un certain recul par rapport à la physique, d'autres peuvent très bien être lues par un étudiant de licence ou de classes préparatoires. Prenons l'exemple du chapitre 5 "fonctions d'une variable ; analyse des signaux". Tout ce qui concerne la première section, à savoir l'information contenue dans un graphe, les dérivées et développements limités, l'intégration et le lien entre la concavité de l'entropie et la thermodynamique relève des premières années de licence ; même l'intégration par la méthode du col ou celle de la phase stationnaire, rarement introduites à ce niveau, peuvent être lues compte tenu des exemples physiques choisis. Le contenu des deuxième section "opérations sur les fonctions et analyse de Dirac" et troisième section "transformée et analyse de Fourier" est en général enseigné en troisième année ; cependant certaines parties comme la transformée de Laplace sont abordables par un étudiant de deuxième année tandis que d'autres comme l'application du filtrage d'un bruit blanc à l'équation de Langevin et aux télécommunications sont d'un niveau master. De même dans la dernière section "optique de Fourier ; filtrage optique", du niveau master, certaines parties peuvent intéresser un étudiant de deuxième année.

Enfin le livre contient de nombreux passages écrits en petits caractères. Dans l'esprit des auteurs ces passages sont considérés comme des compléments réservés à une seconde lecture. C'est le cas par exemple des passages "métrologie" et "ensembles fractaux" du chapitre 1 sur les nombres réels, ou de certaines démonstrations de théorèmes.

Remerciements

Les auteurs tiennent à remercier leurs collègues Y. Gabellini, O. Legrand, F. Peters, L. Petit, J.A. Sepulchre et D. Wilkowski pour la relecture de certains chapitres et leurs suggestions pertinentes, ainsi que G. Gonczi pour son aide à la réalisation de certaines

figures, en particulier le tracé de la surface $S(U, V)$ pour un fluide de Van der Waals. Ils sont également conscients que la rédaction de certaines parties de cet ouvrage a bénéficié indirectement de conversations avec de nombreux autres collègues notamment P. Coullet, J.L. Féménias, J.M. Levy Leblond, J.L. Meunier, ainsi que de l'expérience tirée de nombreux enseignements effectués en licence et maîtrise de physique, licence E.E.A et préparation à l'agrégation de physique.

Nice Mars 2004

Avant propos de la deuxième édition

L'apport principal de cette nouvelle édition concerne le chapitre 11 "Analyse Numérique ; Physique Discrète". Un tel chapitre, inhabituel dans les ouvrages de mathématiques pour la physique, nous semble devoir faire partie de la culture mathématique du physicien en raison de l'importance prise par le calcul et la simulation numériques en physique. Ecrit dans le même esprit, à la fois introductif et synthétique, que les chapitres précédents, il s'efforce de couvrir les méthodes les plus utiles au physicien, en précisant leurs limites (stabilité, précision ...). Le choix des sujets retenus, indispensable compte tenu de l'étendue du domaine, a bénéficié des discussions que nous avons eues avec B. Cessac, T. Corbard, J.L. Féménias, O. Legrand et J. Provost ; qu'ils en soient ici remerciés.

Cette deuxième édition a été aussi l'occasion de quelques ajouts (géométrie sphérique, notions de transport parallèle et de courbure au chapitre 3, équation de Hamilton-Jacobi, formulation variationnelle des équations de Maxwell et géométrie symplectique au chapitre 9) et d'un effort d'amélioration de la lisibilité de l'ouvrage, notamment en faisant ressortir les exemples et en ajoutant quelques figures.

Nice Juillet 2006

Avant propos de la troisième édition

Cette troisième édition a été l'occasion d'introduire le lecteur à deux nouveaux sujets : la physique quantique relativiste au chapitre 4 (équations de Klein-Gordon, de Weyl et de Dirac, champs quantiques, particules et antiparticules) ; la théorie relativiste de la gravitation au chapitre 9 (métrique et courbure de l'espace-temps, équations d'Einstein). Un complément d'analyse numérique (méthode de Béziers utilisée en CAO) a été ajouté au chapitre 11. Enfin l'index a été développé pour permettre un meilleur accès aux mots clé du livre.

Nice Octobre 2010

Chapitre 1

Nombres réels ; grandeurs physiques ; dimensions

Le but de ce chapitre n'est pas de rappeler au lecteur des techniques de calcul qu'il connait, mais d'insister sur des notions non triviales que recouvre l'emploi des réels en physique. Parmi elles figurent la continuité, le caractère algébrique de la plupart des grandeurs physiques, et la notion de dimension qui traduit l'invariance des lois physiques pour certains changements d'unités.

1.1 GRANDEURS PHYSIQUES ; CONTINUITÉ ; PARAMÉTRAGES ADDITIFS

Les nombres servent à comparer entre elles, notamment avec celle qui sert d'unité, des grandeurs de même nature et à transcrire la loi de composition de ces grandeurs. L'exemple qui a servi de prototype depuis la théorie des grandeurs d'Euclide ($\simeq 400$ av. J.-C.) jusqu'à la construction récente des nombres réels (Cantor, Dedekind $\simeq$1870) est la description des longueurs sur une "droite". Aujourd'hui encore, même si la plupart des appareils de mesure avec une aiguille se déplaçant devant un cadran sont remplacés par des appareils à affichage numérique, l'identification ("isomorphisme local") grandeur physique-longueur sur une droite reste valable. En particulier deux notions communes sont très importantes : celle de continuité et celle de paramétrage additif de la loi de composition.

1.1.1 Survol « physique » des ensembles $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$

▷ Composition de grandeurs additives

L'introduction de l'ensemble $\mathbb{N}$ des **nombres entiers** 0,1,2... n'est évidemment pas qu'une question de numérotation. L'important est qu'ils sont générés par une opération récurrente : le passage de n à $n+1$. C'est cette opération fondamentale "addition d'une

unité" que l'on reconnaît dans de très nombreuses situations expérimentales : mise bout à bout de segments identiques sur une droite, répétition d'une rotation autour d'un axe, accrochages successifs de masses pesantes identiques à un ressort, mise en série de forces électromotrices identiques, *etc.* Ces opérations physiques possèdent les propriétés de commutativité et d' associativité qu'on transcrit par l'addition au niveau des nombres. L'entier $q = 1 + 1 + 1 + \cdots + 1$ représente alors la grandeur obtenue en composant q fois avec elle même la grandeur unité, et le nombre 0 l'opération "neutre", qui consiste par exemple à ne rien accrocher au ressort. Enfin si on compose la grandeur associée à q avec elle même, on obtient les grandeurs associées à $2q,\ 3q \ldots$ L'opération de multiplication ainsi introduite est, on le sait, commutative, associative et distributive par rapport à l'addition.

▷ Changements d'unité

On peut adopter un autre point de vue qui est l'analogue du point de vue passif lors d'un changement de référentiel (*cf.* section 3.1.1). Il consiste à considérer que le regroupement q par q de la grandeur unité définit une nouvelle unité q fois plus grande. Alors des grandeurs correspondant aux nombres $x' = 0, 1, 2, 3 \ldots$ lorsqu'elles sont obtenues par composition de cette nouvelle unité correspondent comme on l'a vu aux nombres $x = 0, q, 2q, 3q \ldots$ pour l'ancienne unité. Inversement pour représenter numériquement avec la nouvelle unité des grandeurs correspondant à $x = 0, 1, 2 \ldots p \ldots$ on est amené à introduire les nombres fractionnaires, $x' = \dfrac{p}{q}$ signifiant $x = p$.

▷ Nombres algébriques

Pour de nombreuses grandeurs physiques il existe une **opération inverse** de l'addition, comme par exemple lors de la mise bout à bout de segments orientés. On est alors amené à introduire l'ensemble des **entiers algébriques** $\mathbb{Z} = (\cdots -2, -1, 0, 1, 2 \ldots)$, l'ensemble des **nombres rationnels** $\mathbb{Q} = \left\{ \dfrac{p}{q} \ ;\ p \in \mathbb{Z},\ q \neq 0 \in \mathbb{Z} \right\}$, et les règles du calcul algébrique (les fameuses "règles des signes") afin d'étendre à $\mathbb{Z}$ et $\mathbb{Q}$ les propriétés de l'addition et de la multiplication sur $\mathbb{N}$. Compte tenu de son importance deux sections (1.2.1,2) sont consacrées au **caractère algébrique des grandeurs physiques**.

▷ Commensurabilité et incommensurabilité

Une question a de longue date préoccupé les scientifiques : peut-on toujours comparer deux grandeurs (de même nature) à l'aide de $\mathbb{Q}$? Si oui elles sont commensurables et il existe une unité telle que la première vale $p \in \mathbb{Z}$ et la seconde $q \in \mathbb{Z}$. Mais la réponse à cette question est non, et était déjà connue des Grecs (relations $\dfrac{d^2}{a^2} = 2$ entre diagonale et côté d'un carré, $l = \pi d$ entre circonférence et diamètre d'un cercle...). D'où les autres questions : les grandeurs "irrationnelles" vis-à-vis d'une unité sont-elles une exception ou la règle quasi générale ? Peut-on les comparer entre elles et par quel ensemble doit-on remplacer $\mathbb{Q}$ pour les mesurer ? Les Grecs avaient constaté que les irrationnels sont "plus nombreux" que les rationnels mais ils n'ont pu répondre de manière complète à la deuxième question. On sait aujourd'hui que les rationnels forment un ensemble dense, bien que de mesure de Lebesgue nulle, dans l'ensemble des réels $\mathbb{R}$.

▷ Continuité ; nombres réels

L'approche moderne des nombres réels reprend à son compte toutes les propriétés relatives à $\mathbb{Q}$ (addition, multiplication, ordre, distance...) et donne une méthode de construction par "approximations successives" de l'ensemble $\mathbb{R}$. Cette méthode fait appel à l'**axiome d'Archimède**, qui stipule que quels que soient x et y il existe $n \in \mathbb{N}$ tel que $n|x| \geq |y|$, et à l'**axiome de Cauchy des intervalles emboités**, selon lequel il existe un nombre x et un seul dans la famille d'intervalles $\{I_n = [a_n, b_n]\}$ tels que $I_{n+1} \subset I_n$ et $|b_n - a_n| \to 0$ quand $n \to \infty$. En pratique le premier axiome permet d'encadrer tout réel x par des rationnels $\dfrac{p}{q}$ et $\dfrac{p+1}{q}$, et le second permet par des encadrements emboités $\cdots \left[\dfrac{p_n}{q_n}\left[\dfrac{p_{n+1}}{q_{n+1}} \cdots \dfrac{p_{n+1}+1}{q_{n+1}}\right]\dfrac{p_n+1}{q_n}\right] \cdots$ de définir x comme la limite des $\dfrac{p_n}{q_n}$ ou $\dfrac{p_n+1}{q_n}$ lorsque q_n tend vers l'infini ; on dit que $\mathbb{Q}$ est un **ensemble dense** dans $\mathbb{R}$. Bien évidemment on s'assure que cette définition des réels comme limite de suite de rationnels est unique (notion de suites équivalentes), et qu'elle est compatible avec les opérations définies sur $\mathbb{Q}$ (si $\dfrac{p_n}{q_n}$ tend vers x et $\dfrac{r_n}{s_n}$ vers y alors $\dfrac{p_n r_n}{q_n s_n}$ tend vers $xy = yx$, *etc.*). En base 10, $q_n = 10^n$ ($25,37... = 2\,10^1 + 5\,10^0 + 3\,10^{-1} + 7\,10^{-2} + ...$) ; en base 2, $q_n = 2^n$ (le même nombre est représenté par une suite de 0 et de 1). En toute base on montre que les rationnels correspondent à des suites de chiffres qui deviennent périodiques à partir d'un certain rang, par exemple $\dfrac{1}{12} = 0,83333...$ en base 10 et $0,00\,01\,01\,01...$ en base 2, tandis que les non rationnels correspondent à des suites "chaotiques" (*cf.* π, $\sqrt{2}$, le nombre d'or $\dfrac{1+\sqrt{5}}{2}$ dont la suite des décimales peut servir à générer des nombres aléatoires...).

Le résultat le plus remarquable de cette construction est que $\mathbb{R}$ est un **ensemble complet** : toute suite de Cauchy de nombres réels (suite de nombres x_n tels que $|x_n - x_m| < \epsilon$ arbitrairement petit dès que n et $m > N(\epsilon)$) converge vers un nombre réel. On peut donc parler de "continuité des nombres réels" (à l'inverse de $\mathbb{Q}$ pour lequel les suites de Cauchy "sortent en général de l'ensemble"). Ceci permet par exemple de définir la **continuité d'une fonction** f ou sa **dérivabilité** : f est continue en $x_0 \in \mathbb{R}$ si $|f(x) - f(x_0)|$ peut être rendu arbitrairement petit pour x suffisamment proche de x_0, et est dérivable en x_0 si $\dfrac{f(x_0+\epsilon) - f(x_0)}{\epsilon}$ admet une limite (dérivée en x_0 notée $f'(x_0)$ ou $\dfrac{\mathrm{d}f}{\mathrm{d}x}(x_0)$) pour $\epsilon \to 0$.

▷ Mesure de Lebesgue

Définie par $\mu\big([a,b]\big) = b - a$ pour un intervalle $[a,b]$, elle est nulle pour un point (encadré par les intervalles I_n) de même que pour un ensemble fini ou dénombrable de points (par exemple $\mathbb{Q}$), et elle s'étend à toutes les intersections et unions dénombrables d'intervalles. Elle est invariante par translation et permet de définir l'**intégrabilité** d'une fonction : au sens de **Lebesgue** (resp. **Riemann**) une fonction $y = f(x)$ est intégrable si en partitionnant l'axe y (resp. x) en domaines D_n la quantité $\sum_n y_n\, \mu\big(f^{-1}(D_n)\big)$

(resp $\sum_n f(x_n)\,\mu(D_n)$) a une limite, quels que soient les y_n (resp. x_n) appartenant à D_n, lorsque les $\mu(D_n)$ tendent vers 0 (figure 1a, resp. 1b). Le premier point de vue est mathématiquement plus général (*cf.* $\int_0^1 f(x)\,\mathrm{d}x$, avec $f(x) = 0$ si $x \in \mathbb{Q}$ et 1 ailleurs, qui vaut 1 pour Lebesgue et n'est pas définie pour Riemann) et pas moins physique (*cf.* section 5.3.3).

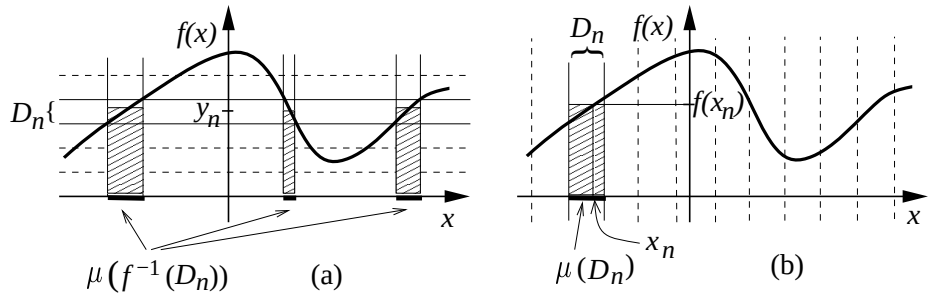


FIGURE 1

1.1.2 Paramétrage additif des lois de composition ; logarithmes

▷ Repérage des grandeurs

Reprenons l'exemple qui consiste à suspendre des masses à un ressort. En pratique, pour les comparer, on "repère" les masses en leur associant le nombre réel x correspondant à l'élongation du ressort. Mais comme la réponse d'un ressort n'est pas linéaire (son comportement n'est approximativement linéaire que pour des petites déformations), l'opération physique qui consiste à accrocher simultanément deux masses ne conduit pas en général à la somme $x_1 + x_2$ des élongations correspondant à chacune des masses mais à un autre nombre $x_1 * x_2$. Cette loi de composition $(x_1, x_2) \to x_1 * x_2$ est cependant commutative, associative et possède un élément neutre que l'on peut toujours noter 0. (Le nombre x peut ne pas avoir d'inverse pour cette loi.)

▷ Paramétrage additif

Pour tout repérage raisonnable ($x_1 * x_2$ continu et dérivable par rapport à x_1 et x_2), on démontre qu'on peut trouver une fonction $\varphi(x)$, définie à une constante multiplicative λ près, telle que :

$$\varphi(x_1 * x_2) = \varphi(x_1) + \varphi(x_2) \quad ; \quad \varphi(0) = 0\ .$$

Donc si on associe à chaque grandeur non plus x_i mais $\varphi_i = \varphi(x_i)$, la loi de composition $x = x_1 * x_2$ devient additive $\varphi = \varphi_1 + \varphi_2$. La **mesure** d'une grandeur, par opposition à son **repérage**, consiste précisément à trouver ce paramétrage additif. La **grandeur unité** est par définition celle pour laquelle $\varphi = 1$. Quant à la liberté du changement de paramétrage $\varphi \to \lambda\varphi$, elle correspond au choix toujours possible d'une unité λ fois plus petite. Remarquons que ce paramétrage additif est la version continue de l'**étalonnage** qui consiste à associer $\varphi = 1$ à un certain nombre x_1, puis $\varphi = \dfrac{1}{q}$ au nombre x_q tel que $\overbrace{x_q * \cdots * x_q}^{q \text{ fois}} = x_1$ et $\varphi = \dfrac{p}{q}$ au nombre $\overbrace{x_q * \cdots * x_q}^{p \text{ fois}}$.

EXEMPLES. Une règle bien graduée traduit additivement la composition "physique" des **translations**. Un cercle bien gradué traduit additivement la composition des **rotations** (le paramètre additif étant l'angle). La loi de composition des **changements de référentiels inertiels** à une dimension, qui s'écrit pour les vitesses relatives

$$v = v_1 * v_2 = (v_1 + v_2)\Big(1 + \frac{v_1 v_2}{c^2}\Big)^{-1}$$

($|v_i| < c$) devient $\varphi = \varphi_1 + \varphi_2$ en introduisant la **rapidité** relative φ définie par :

$$e^{\varphi} = \sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}} \quad \text{ou} \quad \varphi = \frac{1}{2}\ln\frac{1+\frac{v}{c}}{1-\frac{v}{c}} \quad \text{ou} \quad \frac{v}{c} = \frac{e^{\varphi}-e^{-\varphi}}{e^{\varphi}+e^{-\varphi}} = \tanh\varphi\ .$$

(φ est mesurable directement par effet Doppler ; *cf.* section 4.3.3). La vitesse v n'est un paramètre additif que dans la limite galiléenne $v \ll c$ ($\tanh\varphi \simeq \varphi \ll 1$).

▷ Multiplication et logarithme

Le logarithme $\mathrm{Log}_a\, x$ d'un nombre x positif est le paramétrage qui rend additif la loi de multiplication $(x_1, x_2) \to x_1 x_2$ sur $\mathbb{R}^+$:

$$\mathrm{Log}_a\, x_1 x_2 = \mathrm{Log}_a\, x_1 + \mathrm{Log}_a\, x_2 \qquad (\text{ce qui implique } \ \mathrm{Log}_a\, 1 = 0)\ .$$

L'unité du paramétrage, appelée **base a du logarithme**, est le nombre dont le logarithme vaut 1 : $\mathrm{Log}_a\, a = 1$. Comme un changement d'unité, donc ici de base, correspond à l'introduction d'un facteur multiplicatif λ on doit avoir $\mathrm{Log}_a\, x = \lambda \mathrm{Log}_b\, x$; en posant $x = b$ puis $x = a$ dans cette relation on trouve $\lambda = \mathrm{Log}_a\, b = \big(\mathrm{Log}_b\, a\big)^{-1}$. Les logarithmes les plus utilisés sont les logarithmes népérien (à base $e \simeq 2,7$) noté ln et décimal (à base 10) noté log. Il est utile d'avoir en mémoire les deux valeurs approchées $\ln 10 = \big(\log e\big)^{-1} \simeq 2,3$ et $\log 2 \simeq 0,3$. La fonction logarithme vérifie :

$$\frac{\mathrm{d}\,\mathrm{Log}_a\, x}{\mathrm{d}x} = \lim_{\epsilon\to 0}\frac{\mathrm{Log}_a\,(x+\epsilon) - \mathrm{Log}_a\, x}{\epsilon} = \lim_{\epsilon\to 0}\frac{\mathrm{Log}_a\big(1+\frac{\epsilon}{x}\big)}{\epsilon} = \frac{1}{x}\left(\frac{\mathrm{d}\,\mathrm{Log}_a\, x}{\mathrm{d}x}\right)_{x=1}\ .$$

Le **logarithme népérien** est celui pour lequel $\dfrac{\mathrm{d}\,\ln x}{\mathrm{d}x} = \dfrac{1}{x}$ (figure 2b).

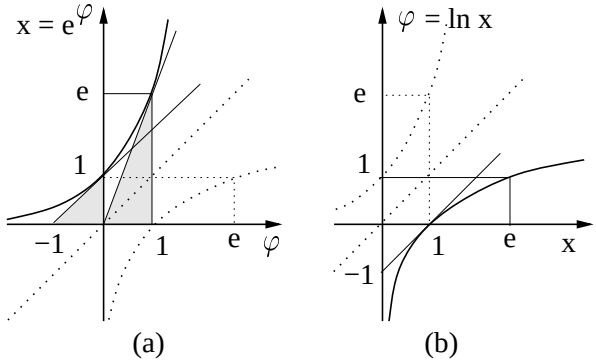


FIGURE 2

EXEMPLES. En chimie, les logarithmes décimaux sont utilisés pour paramétrer des concentrations ou des activités : $pH = -\log\left[H_3O^+\right]$ et $pK_a = -\log\left([A^-][H^+][AH]^{-1}\right)$.
En physique on les introduit pour mesurer des rapports de puissance en **décibels**. Un rapport de 10 dans les puissances correspondant par définition à un écart de 10 décibels (1 Bel) en "échelle logarithmique", on a :

$$\text{“ } \frac{P_2}{P_1} \text{ en décibels ” } = 10 \log \frac{P_2}{P_1} \ .$$

Un rapport de 2 correspond à 3 décibels. Les décibels (dB) sont utilisés en acoustique où les "puissances sonores" sont comparées au seuil d'audibilité P_0 égal à 10^{-12} Wm^{-2}. Ainsi 70 dB correspondent à $10^7 P_0 = 10^{-5}$ Wm^{-2}.
En électronique il servent à mesurer les gains ; par exemple si la tension à la sortie d'un système est V_s et la tension à l'entrée est V_e, sachant que le rapport des puissances est comme $\left(\dfrac{V_s}{V_e}\right)^2$, on pose : "gain en dB" $= 20 \log \dfrac{V_s}{V_e}$. Dans le domaine des fréquences temporelles une **octave** désigne un rapport de 2 et une **décade** un rapport de 10 ; on a donc “ $\dfrac{\omega_2}{\omega_1}$ en octaves (ou décades) ” $= \mathrm{Log}_{2\ (\text{ou } 10)} \dfrac{\omega_2}{\omega_1}$.
En astronomie un rapport de flux lumineux est mesuré en (écart de) **magnitude** par $\Delta m = 2,5 \log \dfrac{F_2}{F_1}$.
Voir aussi la définition de l'entropie à la section 1.1.4.

1.1.3 Fonction et notation exponentielles ; applications

▷ Multiplication et fonction exponentielle

L'opération exponentielle sur $\mathbb{R}$ est l'opération inverse du logarithme (figure 2a) :

$$\varphi = \ln x \Longleftrightarrow x = e^{\varphi} \quad ; \quad \varphi = \mathrm{Log}_a\, x \Longleftrightarrow x = a^{\varphi} \ \ (a > 0) \ .$$

Elle fait donc correspondre à la somme $\varphi_1 + \varphi_2$ d'éléments de $\mathbb{R}$ le produit $x_1 x_2$ d'éléments de $\mathbb{R}^+$. Les propriétés suivantes de l'exponentielle sont des conséquences immédiates de celles du logarithme :

$$a^{b+c} = a^b\, a^c \ \ , \ \ a^b = e^{b \ln a} \ \ , \ \ \left(a^b\right)^c = a^{bc} \ \ , \ \ a^0 = 1 \quad (a > 0) \ .$$

Ainsi par exemple $a^b = e^{b \ln a}$ car $\ln a^b = (\mathrm{Log}_a\, a^b)\ \ln a = b \ln a$.

La fonction e^{φ} (notée aussi $\exp\varphi$) peut être obtenue plus explicitement comme limite ou sous la forme d'un développement en série convergente :

$$e^{\varphi} = \lim_{N\to\infty} \left(1 + \frac{\varphi}{N}\right)^N = 1 + \varphi + \frac{\varphi^2}{2!} + \cdots + \frac{\varphi^p}{p!} + \ldots \quad \left(e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \cdots \simeq 2,7\right) \ .$$

La première expression se déduit du fait que $\ln\left(1 + \dfrac{\varphi}{N}\right)^N = N \ln\left(1 + \dfrac{\varphi}{N}\right)$ tend vers φ lorsque N tend vers l'infini. Quant au développement en série il se déduit de la formule du binôme : $\left(1 + \dfrac{\varphi}{N}\right)^N = \displaystyle\sum_{p=0}^{N} C_N^p \ \left(\frac{\varphi}{N}\right)^p \ \xrightarrow{N\to\infty} \ \sum_{p=0}^{\infty} \frac{\varphi^p}{p!}$. e^{φ} possède la propriété

remarquable d'être sa propre dérivée (*cf.* son développement en série ou la relation $\dfrac{\mathrm{d}e^{\varphi}}{\mathrm{d}\varphi} = \dfrac{\mathrm{d}x}{\mathrm{d}\ln x} = x = e^{\varphi}$).

▷ Applications physiques

En physique les fonctions exponentielles apparaissent souvent à travers leur propriété caractéristique

$$f(\varphi_1+\varphi_2) = f(\varphi_1)\,f(\varphi_2) \quad \text{équivalente à} \quad f'(\varphi) = f'(0)\,f(\varphi)$$

(puisque $\dfrac{f(\varphi+\epsilon)-f(\varphi)}{\epsilon} = f(\varphi)\,\dfrac{f(\epsilon)-1}{\epsilon}$). Par exemple le rapport $p(\tau) = \dfrac{N(t+\tau)}{N(t)}$ des populations de noyaux radioactifs non désintégrés aux instants $t+\tau$ et t doit vérifier

$$p(\tau_1+\tau_2) = p(\tau_1)\,p(\tau_2)$$

si l'on suppose qu'il ne dépend que de l'intervalle de temps τ et pas de t ; en effet :

$$\frac{N(t+\tau_1+\tau_2)}{N(t)} = \frac{N(t+\tau_1+\tau_2)}{N(t+\tau_2)}\;\frac{N(t+\tau_2)}{N(t)}\;.$$

On en déduit que la loi qui décrit la **désintégration radioactive** est du type $p(\tau) = \exp(-\lambda\tau)$, ou $\exp\left(-\dfrac{\tau\,\ln 2}{T}\right)$ en introduisant la **demi-période** T c'est-à-dire le temps au bout duquel la moitié des noyaux de l'échantillon se sont désintégrés. On établit par le même raisonnement que l'atténuation de l'intensité d'une onde se propageant selon Ox dans un milieu homogène est donnée par la **loi d'absorption de Beer**

$$I(x+l) = I(x)\,e^{-\alpha l}$$

ou $\alpha l = \ln\dfrac{I(x)}{I(x+l)} \simeq 2,3\,\log\dfrac{I(x)}{I(x+l)}$. Par exemple une atténuation de 0,2 dB par kilomètre correspond à $\log\dfrac{I(x)}{I(x+l)} = 2\,10^{-2}$ et $l = 10^3\mathrm{m}$, soit $\alpha \simeq 4,6\times 10^{-5}\mathrm{m}^{-1}$.

▷ Développement de Taylor et écriture exponentielle

L'approche de l'exponentielle e^{φ} par une limite ou une série ne fait appel qu'aux propriétés des deux opérations $+$ et $\times$ sur les réels. L'extension de l'exponentielle à d'autres objets mathématiques que les réels (nombres complexes, matrices carrées, opérateurs différentiels...) est très utilisée. Ainsi, à titre d'exemple la formule de Taylor (pour les "bonnes" fonctions)

$$f(x+a) = f(x) + a\,f'(x) + \frac{a^2}{2!}\,f''(x) + \cdots + \frac{a^p}{p!}\,f^{(p)}(x) + \cdots = e^{a\frac{\mathrm{d}}{\mathrm{d}x}}\,f(x)$$

s'écrit à l'aide de l'exponentielle de $\dfrac{\mathrm{d}}{\mathrm{d}x}$. Cela se comprend en remarquant qu'on passe de $f(x)$ à $f\big(x+\dfrac{a}{N}\big)$ par $f\big(x+\dfrac{a}{N}\big) \simeq f(x)+\dfrac{a}{N}\,f'(x) = \left(1+\dfrac{a}{N}\,\dfrac{\mathrm{d}}{\mathrm{d}x}\right)f(x)$, et par conséquent de $f(x)$ à $f(x+a)$ par l'application de $\lim\limits_{N\to\infty}\left(1+\dfrac{a}{N}\,\dfrac{\mathrm{d}}{\mathrm{d}x}\right)^N = e^{a\frac{\mathrm{d}}{\mathrm{d}x}}$. (On établit de même pour l'opération de dilatation : $f(e^a x) = e^{ax\frac{\mathrm{d}}{\mathrm{d}x}}\,f(x)$.) On dit que l'opérateur dérivation

$\dfrac{\mathrm{d}}{\mathrm{d}x}$ est le générateur des translations $f(x) \to f(x+a)$. En **physique quantique** les générateurs des transformations "continues" effectuées sur les états ou fonctions d'onde, translation d'espace, de temps, rotation, sont associés à des observables, quantité de mouvement, énergie, moment cinétique (*cf.* sections 4.4).

▷ Fonctions hyperboliques (figure 3)

A la fonction e^{φ} sont reliées les fonctions cosinus hyperbolique (paire), sinus et tangente hyperboliques (impaires)

$$\cosh\varphi = \frac{e^{\varphi}+e^{-\varphi}}{2} \quad , \quad \sinh\varphi = \frac{e^{\varphi}-e^{-\varphi}}{2} \quad , \quad \tanh\varphi = \frac{\sinh\varphi}{\cosh\varphi} \, ,$$

dont les propriétés principales (analogues à celles des fonctions trigonométriques) sont des conséquences immédiates de celles de e^{φ} :

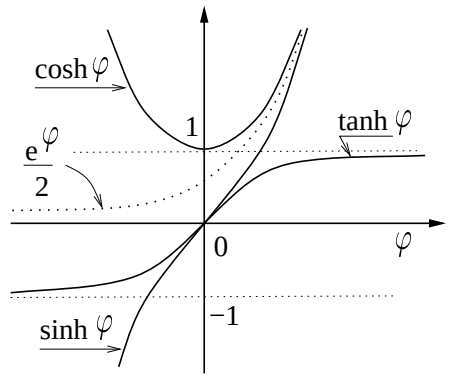


FIGURE 3

$$\cosh\varphi = \sum_0^{\infty} \frac{\varphi^{2n}}{(2n)!} \quad ; \quad \sinh\varphi = \sum_0^{\infty} \frac{\varphi^{2n+1}}{(2n+1)!} \quad ; \quad \tanh\varphi = \varphi - \frac{\varphi^3}{3} + \cdots \; ;$$

$$(\cosh\varphi)' = \sinh\varphi \quad ; \quad (\sinh\varphi)' = \cosh\varphi \quad ; \quad (\tanh\varphi)' = (\cosh\varphi)^{-2} \; ;$$

$$\cosh^2\varphi - \sinh^2\varphi = 1 \quad ; \quad \cosh(\varphi_1+\varphi_2) = \cosh\varphi_1\cosh\varphi_2 + \sinh\varphi_1\sinh\varphi_2 \; ;$$

$$\sinh(\varphi_1+\varphi_2) = \sinh\varphi_1\cosh\varphi_2 + \cosh\varphi_1\sinh\varphi_2 \; ;$$

$$\tanh(\varphi_1+\varphi_2) = \frac{\tanh\varphi_1+\tanh\varphi_2}{1+\tanh\varphi_1\tanh\varphi_2} \; .$$

On les rencontre essentiellement en relativité d'Einstein et en mécanique statistique.

1.1.4 Mesure additive du désordre microscopique ; grands nombres et entropie

▷ Définition et exemples

L'entropie S introduite par Clausius en 1854 reste encore pour certains une grandeur mystérieuse. En fait ses propriétés découlent de l'interprétation qu'en a donné Boltzmann ($\sim$1877) : S est une mesure additive du "désordre microscopique" (ou plutôt de la "liberté

microscopique", voir ci-dessous) donnée par

$$S = k_B \ln \Omega \qquad \left(k_B = \frac{R}{\mathcal{N}} = 1,4 \times 10^{-23}\,\mathrm{J.K^{-1}}\right) .$$

Ω est le nombre d'états microscopiques (spécifiés chacun par les positions, quantités de mouvement et états quantiques de toutes les particules) compatibles avec les valeurs V, U, *etc.* des grandeurs volume total, énergie interne totale, *etc.* qui caractérisent un état d'équilibre macroscopique.

EXEMPLES. Pour un **gaz parfait** constitué de N particules dans un volume V, le volume "libre" moyen par particule est $\dfrac{V}{N}$ et le nombre de configurations de position est donc proportionnel à $\left(\dfrac{V}{N}\right)^N$. Pour un **gaz parfait monoatomique**, on établit (*cf.* section 7.2.1.) que la contribution à Ω du désordre dans l'espace des vitesses ou quantités de mouvement est proportionnelle à $\big(U/N\big)^{\frac{3}{2}N}$; on peut dire aussi que $(2mU/N)^{\frac{3}{2}}$ est un volume "libre" typique dans l'espace (p_x, p_y, p_z) pour une particule d'énergie $\dfrac{p^2}{2m} \simeq \dfrac{U}{N}$.

D'autres exemples (justifiables aussi par des arguments qualitatifs ou dimensionnels) sont : $\Omega \propto \big(\frac{U}{N}\big)^{\frac{3}{2}N}\big(\frac{U}{N}\big)^{\frac{3}{2}N}$ pour un **solide** dont chacun des N atomes est assimilé à un oscillateur à trois dimensions (dans Ω un facteur est lié à la vitesse et l'autre à l'amplitude de la vibration ; V n'apparaît pas car il n'y a pas de "liberté" de position dans un solide) ; $\Omega \propto \max_N \big(\frac{V}{N}\big)^N\big(\frac{U}{N}\big)^{3N}$ d'où $S \propto V^{\frac{1}{4}}U^{\frac{3}{4}}$ pour le **corps noir** (ici $\frac{U}{N} \simeq cp$ et on prend le $\max_N$ car le nombre de photons n'est pas fixé à priori ; *cf.* remarque ci-dessous et sections 1.3.2 et 7.4.1).

▷ Conséquences

Même si Ω n'est calculable que dans un petit nombre de cas, la définition de S permet de comprendre les bases de la thermodynamique. On en déduit que $S = 0$ ($\Omega = 1$) pour un **système mécanique** macroscopique (par exemple un piston, caractérisé par une position et une vitesse), et que S est une grandeur extensive

$$S = S_1 + S_2 \quad (\Omega = \Omega_1\Omega_2)$$

pour un système composé de deux corps isolés l'un de l'autre par des contraintes. On comprend aussi l'**évolution irréversible vers l'équilibre** : si des contraintes sont supprimées (système plus "libre") alors Ω_f, qui correspond à l'état macroscopique final moins contraint, est non seulement plus grand que Ω_i (correspondant à l'état initial) mais en fait "infiniment plus grand". Par exemple si on offre à un gaz un volume supplémentaire ΔV initialement vide (détente de Joule à U constant) on a, même pour des valeurs aussi petites que $N = 10^{-9}$ moles et $\dfrac{\Delta V}{V} = 10^{-6}$:

$$\frac{\Omega_f}{\Omega_i} = \left(\frac{V + \Delta V}{V}\right)^N \simeq \exp\!\left(N\frac{\Delta V}{V}\right) = \exp\left(6 \times 10^8\right)!$$

De même si deux gaz monoatomiques ayant les mêmes valeurs fixées de N et V mais des énergies $U_1 \neq U_2$ sont mis en contact thermique, on a

$$\frac{\Omega_f}{\Omega_i} = \left(\frac{\left(\frac{U_1+U_2}{2}\right)^2}{U_1U_2}\right)^{\frac{3}{2}N} = \left(1 + \frac{(U_2-U_1)^2}{4U_1U_2}\right)^{\frac{3}{2}N} ,$$

rapport qui dans la pratique est lui aussi quasi infini. Si on admet que tous les états microscopiques sont équiprobables, on en déduit que le système des deux corps va, sous l'effet des interactions désormais permises par la suppression des contraintes, évoluer de façon "certaine" vers l'état macroscopique correspondant au maximum de Ω donc de S. On notera que $\dfrac{\Omega_f}{\Omega_i} = \exp\dfrac{S_f - S_i}{k_B} = \exp(10^8)$ correspond à un $\Delta S = 1,4 \times 10^{-15}\,\mathrm{J.K^{-1}}$ extrêmement petit à l'échelle macroscopique.

REMARQUES. 1) Ω_f calculé ci-dessus pour les deux gaz correspond à $\Omega_f = \max_{(1,2)} \Omega_1\Omega_2$ (ou $S_f = \max_{(1,2)}(S_1 + S_2)$) avec $U_1 + U_2 = U$ fixé. Pour le calcul de S il revient au même de prendre $\Omega_f = \Sigma_{(1,2)}\,\Omega_1\Omega_2$ ($= \Omega$ total) car dans une somme $\Omega = \Sigma_{i=1}^{p}\,\lambda_i^N$ on voit, en mettant en facteur le terme λ_{max}^N, que c'est lui qui donne $\ln\Omega$ pour N grand (*cf.* aussi section 5.1.2). 2) La **loi de Boltzmann** est elle aussi une conséquence directe de $S = k_B\,\ln\Omega$ (*cf.* section 10.2.3).

1.2 CARACTÈRE ALGÉBRIQUE DES GRANDEURS PHYSIQUES

1.2.1 Pensée « naïve » et pensée algébrique

Le langage usuel fait toujours référence à des quantités positives : distance de déplacement sur une route dans une direction ou son opposée, avance ou retard à un rendez vous, avoir ou dette dans un compte, *etc.* Il en est de même des variations de grandeurs échangées : gain ou perte d'argent, *etc.* Ce mode de pensée "naïf" revient à raisonner comme si on ne connaissait que les nombres positifs.

▷ Discours algébrique

Reprenons les exemples ci-dessus sous l'angle algébrique. A un déplacement sur une droite orientée (axe), on associe un nombre x positif si le déplacement se fait dans le sens de l'axe et un nombre négatif dans le cas contraire. L'intérêt de ce point de vue est que la loi de composition des translations correspond toujours à une addition $x = x_1 + x_2 + x_3 \cdots$ de nombres algébriques, au lieu d'une suite d'additions et de soustractions de nombres positifs dépendant du cas particulier considéré ; pour le choix d'orientation opposée, les x_i deviennent $-x_i$, mais leur loi de composition demeure inchangée. En ce qui concerne le **temps**, celui-ci est traditionnellement orienté du passé vers le futur, par référence aux phénomènes irréversibles de la vie courante. Il n'empêche que les notions d'avance et de retard peuvent être considérées comme algébriques, une avance de $\tau = -1$ heure correspondant à un retard de $-\tau = 1$ heure (et réciproquement). Enfin on sait bien que dans un compte on peut considérer une dette comme un avoir négatif et faire le bilan en ajoutant les avoirs algébriques. De même dans un échange il est équivalent de donner 100 euros à un ami (gain algébrique pour lui de $+100$) ou de le libérer d'une dette de même valeur (gain algébrique pour lui de $-(-100)$). Cette équivalence lors d'échange de grandeurs algébriques conservées apparaît en physique par exemple à propos des **courants électriques**, associés à un flux dans un sens ou son opposé de charges positives ou négatives, et de la notion de **pression cinétique**, associée à un flux dans un sens ou son opposé de quantités de mouvement ; sur la figure 4, dans les deux cas, la région 2 gagne une charge q et une quantité de mouvement $m\vec{v}$, et la région 1 gagne $-q$ et $-m\vec{v}$.

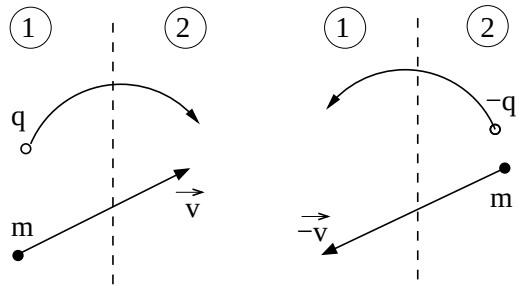


FIGURE 4

▷ Caractère algébrique des grandeurs physiques

Ce caractère tient souvent au fait que la grandeur considérée est la composante d'une grandeur vectorielle sur un axe (produit scalaire du vecteur avec le vecteur unitaire de l'axe) : circulation d'un champ vectoriel le long d'une courbe orientée, flux d'un champ vectoriel à travers une surface orientée, *etc.* Pour de telles grandeurs un changement du sens d'orientation entraîne un changement de signe. Mais il existe aussi des grandeurs scalaires algébriques (charge électrique, tension d'un ressort, *etc.*). Pour ces dernières, comme pour le temps, la convention sur leur signe fait partie de la définition de la grandeur.

▷ Conséquences sur les lois physiques

Les lois physiques $v = Ri$, $e = -\dfrac{\mathrm{d}\varphi}{\mathrm{d}t}$, *etc.* font intervenir des grandeurs algébriques v, i, e, $\varphi \ldots$ Leur écriture dépend donc des **conventions de signe** adoptées pour ces grandeurs, conventions qu'il est important de préciser et de retenir. Une fois précisées les conventions relatives à des grandeurs A et B, on peut se rappeler si une loi s'écrit $A = B$ ou $A = -B$ en la testant dans une situation particulière. Si par exemple pour $A > 0$ la "physique dit" (l'expérience montre) que B doit être positif (resp. négatif), la loi est $A = B$ (resp. $A = -B$) ; en vertu de son caractère algébrique la formule ainsi obtenue sera toujours valable. On verra sur des exemples qu'il n'y a pas lieu de s'inquiéter de ces multiples écritures : les résultats physiques ne dépendent pas des conventions.

EXEMPLE. **Ressort** (figure 5).

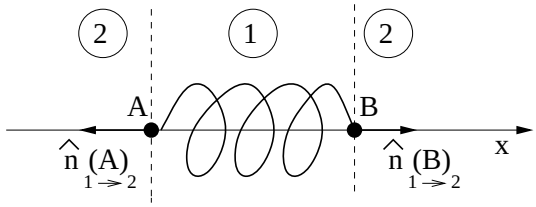


FIGURE 5

L'**allongement** δl d'un ressort est algébrique ($\delta l < 0$ dans le cas d'une compression). Sa **tension** T, qui caractérise l'état mécanique du ressort, l'est aussi. On la définit en introduisant les forces $\vec{F}_{1\to 2}$ exercées par le ressort (système 1) sur l'extérieur (système 2) : $\vec{F}_{1\to 2} = -T\,\hat{n}_{1\to 2}$, $\hat{n}_{1\to 2}$ désignant le vecteur unitaire dirigé vers l'extérieur. La relation bien connue $T = K\,\delta l$ avec $K > 0$ est elle aussi algébrique : $T > 0$ si $\delta l > 0$ (ressort tendu) et $T < 0$ si $\delta l < 0$ (ressort comprimé). Jusqu'à présent on n'a pas choisi d'orientation sur la direction du ressort. Cela devient indispensable pour étudier le mouvement de masses m_A et m_B accrochées à ses extrémités A et B. Avec une orientation de A vers B et en notant x_A et x_B les écarts algébriques par rapport à la position d'équilibre on a $\delta l = x_B - x_A$, $m_A\ddot{x}_A = T$ et $m_B\ddot{x}_B = -T$. Un choix opposé (de B vers A) conduit à des relations différentes $\delta l = x_A - x_B$, $m_A\ddot{x}_A = -T$ et $m_B\ddot{x}_B = T$, mais il ne modifie pas les équations du mouvement $m_A\ddot{x}_A = -K(x_A - x_B)$ et $m_B\ddot{x}_B = -K(x_B - x_A)$ obtenues en remplaçant T par $K\,\delta l$.

1.2.2 Conventions et lois de l'électricité

▷ Charges et courants

Une convention historique, un peu malheureuse, a attribué une charge électrique positive à une barre de verre frottée avec une peau de chat. Comme on le sait aujourd'hui la barre perd alors des électrons et cela a conduit à attribuer une **charge négative aux électrons**. Une intensité électrique étant un flux de charges à travers une surface, on doit préciser l'orientation de la normale à cette surface, disons de 1 vers 2 (*cf.* figure 6). Un transfert de charges positives de 1 vers 2, ou de charges négatives de 2 vers 1, correspond à une intensité i de courant positive. (Plus savamment $i = \iint \vec{j}\cdot\overrightarrow{\mathrm{d}S}$ avec $\vec{j} = nq\vec{v}$ où n

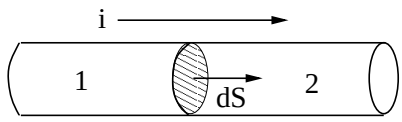


FIGURE 6

est la densité des charges q animées d'une vitesse moyenne $\vec{v}$). Rappelons aussi que pour la **charge q d'un condensateur**, il est nécessaire de préciser l'armature qui porte cette charge, l'autre portant la charge $-q$ au même instant.

▷ Tensions et forces électromotrices

La convention sur les charges étant fixée, la valeur algébrique du **potentiel** V_A en A est déterminée par le fait que qV_A est l'énergie potentielle d'une charge q placée en A (l'origine des potentiels étant prise "à l'infini" ou "à la masse"). Quant à la **tension** v aux bornes d'un dipole AB, elle est égale à la différence des potentiels aux bornes, et vaut soit $v = V_A - V_B$ soit $v = V_B - V_A$ suivant la convention choisie qui doit donc être précisée ; on rappelle que $V_A - V_B = \int_A^B \overrightarrow{E}_{es}\cdot\overrightarrow{\mathrm{d}l}$ est la circulation du **champ électrostatique** $\overrightarrow{E}_{es} = -\overrightarrow{\mathrm{grad}V}$. Les **forces électromotrices d'induction** sont définies pour une portion de circuit orienté par $e_{AB} = \int_A^B \overrightarrow{E}_{em}\cdot\overrightarrow{\mathrm{d}l}$ (circulation du **champ électrique induit** $\overrightarrow{E}_{em}$)

et pour un circuit fermé orienté par $e = \oint_C \overrightarrow{E}_{em} \cdot \overrightarrow{\mathrm{d}l}$; leur signe dépend donc du choix d'orientation sur le circuit (sens de $\overrightarrow{\mathrm{d}l}$). De même **la loi de l'induction** $e = -\dfrac{\mathrm{d}\varphi}{\mathrm{d}t}$ $(= \oint_C \overrightarrow{E}_{em} \cdot \overrightarrow{\mathrm{d}l})$, avec $\varphi = \iint_S \overrightarrow{B} \cdot \overrightarrow{\mathrm{d}S}$, suppose qu'il existe la relation "du tire-bouchon" (*cf.* section 3.1.3) entre le sens de parcours sur C et l'orientation de la normale à la surface S qui s'appuie sur C. (Pour une inductance $\varphi = Li$ avec $L > 0$.)

▷ Dipoles (figure 7)

Les lois des dipoles dépendent de deux conventions de signe qui sont précisées sur les schéma par des flèches : de façon standard la flèche associée à i est aussi celle qui oriente le circuit, tandis que celle associée à v précise que v est égale au potentiel à l'extrémité de la flèche moins le potentiel à son origine. Ainsi pour les figures 7a (en présence de champ électromoteur) 7b et 7c, on a $v = V_A - V_B$ et les lois s'écrivent :

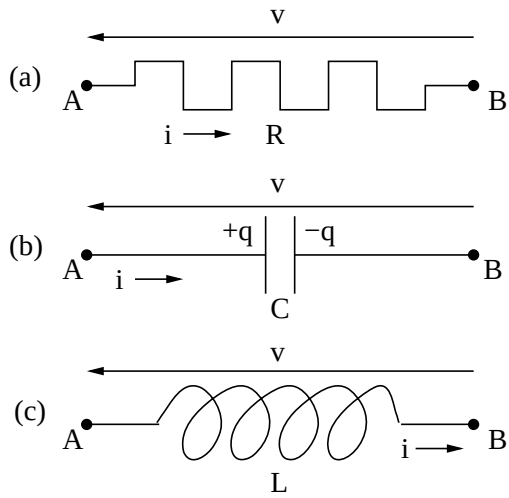


FIGURE 7

$$v = Ri - e_{AB} \qquad ; \qquad v = \frac{q}{C} \; ;$$

$$v = L\frac{\mathrm{d}i}{\mathrm{d}t} \; (= -e_{AB}) \qquad ; \qquad i = \frac{\mathrm{d}q}{\mathrm{d}t} \; .$$

$(Ri = \int_A^B (\overrightarrow{E}_{es} + \overrightarrow{E}_{em}) \cdot \overrightarrow{\mathrm{d}l}$ est la loi d'Ohm). Le choix du sens de la flèche de v opposé à celui de i s'appelle **convention récepteur** (voir justification plus loin). La convention opposée, flèches de v et i de même sens, s'appelle **convention générateur** ; son choix aurait conduit pour le premier dipole par exemple à la loi $v\ (= V_B - V_A) = e_{AB} - Ri$.

▷ Réseaux

En électricité la relation $\oint_C \overrightarrow{E}_{es} \cdot \overrightarrow{\mathrm{d}l} = 0$ est connue sous le nom de **loi des mailles** tandis que la relation de conservation de la charge $\oiint_S \vec{j} \cdot \overrightarrow{\mathrm{d}S} = 0$ l'est sous le nom de

loi des noeuds (S surface fermée entourant le noeud). Ces lois s'écrivent aussi :

$$\epsilon_1 v_1 + \epsilon_2 v_2 + \cdots + \epsilon_n v_n = 0 \qquad \text{et} \qquad \epsilon'_1 i_1 + \epsilon'_2 i_2 + \cdots + \epsilon'_p i_p = 0 \ .$$

Les v_j sont les tensions aux bornes des n dipoles qui constituent la maille (circuit fermé C) et chaque ϵ_j vaut ± 1 suivant que la flèche correspondant à v_j est dans le même sens ou en sens inverse du sens de parcours choisi sur C ; pour les courants i_k dans les p branches qui constituent un noeud chaque ϵ'_k vaut ± 1 suivant que la flèche associée à i_k sort du noeud ou se dirige vers lui. L'introduction des ϵ et des ϵ' est indipensable car lorsqu'il y a plusieurs mailles il est impossible de choisir les conventions sur les tensions et les courants de sorte qu'il valent tous $+1$.

▷ Système électromécanique

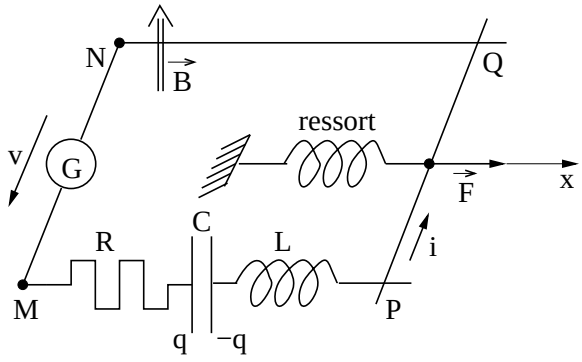


FIGURE 8

EXEMPLE (figure 8). Sur la figure sont précisées les conventions de signes choisies (parmi 64 possibles !) pour le déplacement x de la barre $\overrightarrow{PQ} = \vec{l}$ de masse m, pour la force extérieure $\overrightarrow{F}$ a priori variable, pour le champ magnétique $\overrightarrow{B}$ constant et perpendiculaire au plan du circuit, pour le courant i, la charge q du condensateur et la tension $v = V_M - V_N$ aux bornes du générateur G. Par projection des forces $-K\vec{r}$, $-f\dot{\vec{r}}$, $\overrightarrow{F}$ et $i\vec{l} \wedge \overrightarrow{B}$ sur l'axe x on obtient l'équation du mouvement $m\ddot{x} = -Kx - f\dot{x} + F + ilB$. Si l'orientation de cet axe est changée, mais pas les autres conventions, il faut changer $F + ilB$ en $-F - ilB$ ou multiplier x, $\dot{x}$ et $\ddot{x}$ par -1. En introduisant les symboles ϵ_x, ϵ_F, ϵ_i et ϵ_B valant 1 pour la convention de départ sur x, $\overrightarrow{F}$, i et $\overrightarrow{B}$, et -1 pour la convention opposée, on écrit l'équation de la mécanique sous la forme générale (qui permet de voir l'influence du choix effectué) :

$$\epsilon_x(m\ddot{x} + f\dot{x} + Kx) = \epsilon_F F + l\,(\epsilon_i i)(\epsilon_B B) \ .$$

De même les équations de l'électricité, qui s'écrivent $V_M - V_N = v = Ri - e + L\dfrac{\mathrm{d}i}{\mathrm{d}t} + \dfrac{q}{C}$, $i = \dfrac{\mathrm{d}q}{\mathrm{d}t}$ et $e = -\dfrac{\mathrm{d}\varphi}{\mathrm{d}t} = -Bl\dot{x}$ avec la convention initiale, s'écrivent de façon générale $V_M - V_N = \epsilon_v v = \epsilon_i\Big(Ri - e + L\dfrac{\mathrm{d}i}{\mathrm{d}t}\Big) + \epsilon_q \dfrac{q}{C}$, $\epsilon_i i = \epsilon_q \dfrac{\mathrm{d}q}{\mathrm{d}t}$ et $\epsilon_i e = -l\,(\epsilon_B B)(\epsilon_x \dot{x})$.

En éliminant i entre ces équations on obtient le bilan électrique :

$$V_M - V_N = \epsilon_v v = \epsilon_q \Big(\frac{q}{C} + R\frac{\mathrm{d}q}{\mathrm{d}t} + L\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} \Big) + l\,(\epsilon_B B)(\epsilon_x \dot{x}) \ .$$

On constate que ϵ_i a disparu avec i et que les termes en q, $\dot{q}$ et $\ddot{q}$ apparaissent avec le même signe ϵ_q en facteur comme attendu physiquement (*cf.* décharge d'un condensateur dans une résistance ou oscillations d'un circuit LC). On obtient les bilans énergétiques mécanique et électrique en multipliant les équations mécanique et électrique respectivement par $\epsilon_x \dot{x}$ et $\epsilon_q \dot{q}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}m\dot{x}^2 + \frac{1}{2}Kx^2\Big) = -f\dot{x}^2 + \epsilon_F\epsilon_x F\dot{x} + \epsilon_i\epsilon_B\epsilon_x liB\dot{x}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}Li^2 + \frac{1}{2}\frac{q^2}{C}\Big) = -Ri^2 + \epsilon_v\epsilon_i vi - \epsilon_i\epsilon_B\epsilon_x liB\dot{x} \ .$$

Dans le bilan global obtenu par addition "$\frac{\mathrm{d}}{\mathrm{d}t}$(énergie mécanique + énergie électrique) = puissances (fournie + dissipée)", on remarque que quelles que soient les conventions le terme magnétique disparait et que la puissance dissipée $-f\dot{x}^2 - Ri^2$ est négative. Au contraire l'expression (mais pas la valeur) de la puissance fournie $\epsilon_v\epsilon_i vi + \epsilon_F\epsilon_x F\dot{x}$ dépend des conventions. La puissance électrique fournie s'écrit $+vi$ pour la convention générateur (d'où son nom) et $-vi$ pour la convention récepteur ; quant à la puissance mécanique elle s'écrit $F\dot{x}$ lorsque les conventions pour l'axe x et $\overrightarrow{F}$ sont les mêmes (cas habituel) et $-F\dot{x}$ dans le cas contraire.

1.3 GRANDEURS PHYSIQUES ; DIMENSIONS

Le calcul aux dimensions introduit par Reynolds (1883) est extrêmement précieux en physique. Il permet de vérifier qu'une expression n'est pas absurde, de deviner ou de retrouver (au moins partiellement) une loi connaissant les grandeurs susceptibles d'être mises en jeu, d'obtenir des ordres de grandeur, de simplifier des équations en remplaçant par 1 les paramètres dont le rôle est "trivial", *etc.* Il est fortement chargé de sens, physique comme mathématique, car il correspond à une invariance remarquable des lois physiques, celle relative aux changements d'unités.

1.3.1 Changements d'unités et invariance des lois physiques

▷ Grandeurs, nombres et dimensions

Soient des grandeurs (longueur, surface, temps, vitesse, intensité électrique, résistance...) mesurées par les **nombres** $g_1, g_2 \cdots$ lorsqu'on a choisi pour chacune d'elle une grandeur unité. Une loi de la physique mettant en jeu ces grandeurs s'exprime par une relation $f(g_1, g_2 \cdots) = 0$. Dans des changements d'unités arbitraires ces nombres deviennent $g'_1 = \lambda_1 g_1$, $g'_2 = \lambda_2 g_2 \cdots$ et a priori on n'a plus $f(g'_1, g'_2 \cdots) = 0$. Par contre on constate expérimentalement que l'équivalence

$$f(g_1, g_2 \cdots) = 0 \iff f(\lambda_1 g_1, \lambda_2 g_2 \cdots) = 0$$

est satisfaite si les changements d'unités sont convenablement reliés entre eux : **invariance de la loi par changements d'unités**. C'est cette relation entre changements d'unités que la notion de dimension sert à préciser, et cette invariance qui fait que les physiciens écrivent les lois $f(G_1, G_2, \dots) = 0$ avec les grandeurs physiques G_i plutôt qu'avec les nombres g_i (de même qu'on écrit des relations entre vecteurs plutôt qu'entre leurs composantes).
Une observation importante de la physique classique galiléenne est que toutes les lois sont laissées invariantes par des changements d'unités "non corrélés" des quatre grandeurs "de base" (la température mise à part) **masse, longueur, temps et charge électrique**, les changements d'unités des autres grandeurs étant eux définis à partir des précédents. Autrement dit la dimension d'une grandeur quelconque s'exprime à partir des dimensions des quatre grandeurs de base. En particulier une grandeur est **sans dimension** si il est impossible de changer son unité sans modifier l'écriture des lois dans lesquelles elle intervient, et l'invariance ci-dessus entraîne que les égalités en physique ne peuvent avoir lieu qu'entre grandeurs de même dimension.

EXEMPLES GÉOMÉTRIQUES. La relation $S = a^2$ entre l'**aire** d'un carré et son côté est valable en géométrie euclidienne si les unités de longueur et de surface sont par exemple le mètre et le mètre carré. Si on divise ces deux unités par 10, alors $S' = 10S$, $a' = 10a$ et la relation entre surface et côté devenue $S' = 10^{-1}a'^2$ est modifiée. Pour que la loi $S = a^2$ reste invariante, il faut que les changements d'unité $a' = \lambda_L a$ et $S' = \lambda_S S$ soient reliés par $\lambda_S = \lambda_L^2$. On dit que la dimension d'une surface (d'un volume) est le carré (le cube) de celle d'une longueur et on écrit symboliquement $[S] = [L]^2$, $[V] = [L]^3$ (en oubliant parfois les crochets). Pour les lois mettant en jeu un **angle** comme $l = a\theta$ ou $S = \dfrac{1}{2}a^2\theta$ (longueur d'un arc et surface d'un secteur circulaires de rayon a, θ en radians), on constate que leur invariance implique $\lambda_\theta = 1$; l'angle est donc une grandeur sans dimension ($[\theta] = 1$), mais pas sans unité ! Remarquons que ce caractère "sans dimension" n'a rien à voir avec le choix de l'unité radian ; si l'unité correspond à un tour, on a $l = 2\pi a\theta$ et $S = \pi a^2\theta$, formules qui elles aussi ne sont invariantes que si $\lambda_\theta = 1$.
APPLICATION : démonstration "dimensionnelle" du **théorème de Pythagore**. Un triangle rectangle étant entièrement déterminé par son hypothénuse a et un angle aigu α sa surface est nécessairement de la forme $a^2\, f(\alpha)$; en considérant les deux "petits" triangles rectangles d'hypothénuses b et c obtenus en abaissant la hauteur issue de l'angle droit du "grand" triangle, on obtient $a^2\, f(\alpha) = b^2\, f(\alpha) + c^2\, f(\alpha)$.
REMARQUE. Il faut noter que le caractère euclidien de l'espace est essentiel pour qu'une longueur ait une dimension. Si sur la sphère on considère par exemple un triangle rectangle, formé de trois arcs de grand cercle (géodésiques) dont deux se coupent à angle droit, la relation entre hypothénuse a, côté b et angle opposé $\widehat{B}$ s'écrit, pour des unités de longueur et d'angle bien choisies, $\sin b = \sin a \sin \widehat{B}$ (au lieu de $b = a \sin \widehat{B}$ dans le cas euclidien). On en conclut qu'en **géométrie sphérique** longueur et angle sont sans dimension car tout changement d'unités pour a et b entraîne une modification de l'écriture de la relation contrairement au cas euclidien.

EXEMPLES PHYSIQUES. L'invariance de la loi de transformation de Galilée des intervalles d'espace-temps $x' = x - Vt$, $t' = t$ implique $\lambda_V = \lambda_L\, \lambda_T^{-1}$ donc

$[V] = [L][T]^{-1}$, celle de $\overrightarrow{F} = m\dfrac{\mathrm{d}\vec{v}}{\mathrm{d}t} = q(\overrightarrow{E} + \vec{v} \wedge \overrightarrow{B})$ implique $[F] = [M][L][T]^{-2}$, $[\overrightarrow{E}] = [\vec{v}][\overrightarrow{B}] = [\overrightarrow{F}][Q]^{-1}$, *etc.* Souvent, pour des raisons de commodité, on utilise l'**énergie**, de dimension $[E] = [M][L]^2[T]^{-2}$, à la place de la masse pour le calcul aux dimensions. En effet on la retrouve dans de nombreuses expressions qui touchent tous les domaines de la physique telles que $\dfrac{1}{2}mv^2$ ou mc^2 en mécanique, qV et $\dfrac{q^2}{4\pi\epsilon_0 r}$ en électrostatique, $\dfrac{-Gm^2}{r}$ en gravitation, $h\nu$ en quantique, *etc.* On en déduit par exemple $[V] = [E][Q]^{-1}$, $[\epsilon_0]^{-1} = [E][L][Q]^{-2} = [\mu_0][L]^2[T]^{-2}$ (car $\epsilon_0\mu_0 c^2 = 1$ sans dimension), $[G] = [E][L][M]^{-2} = [E]^{-1}[L]^5[T]^{-4}$, $[h] = [E][T]$ et $[\alpha] = 1$ ($\alpha = \dfrac{q_e^2}{4\pi\epsilon_0\hbar c}$ constante de structure fine).

Les paragraphes qui suivent montrent que la notion de dimension d'une grandeur physique n'est pas absolue.

▷ Dimensions et métrologie

Il ne faut pas confondre calcul aux dimensions et métrologie. Le but de cette dernière est de choisir les **étalons** les plus fiables et précis possibles pour certaines grandeurs, et tels que les mesures des autres grandeurs puissent s'y ramener avec une précision supérieure à celle que l'on aurait avec des étalons propres à ces autres grandeurs. Le nombre des étalons n'a pas toujours été le même dans l'histoire de la physique. C'est ainsi qu'on a rattaché la calorie au Joule en fixant l'équivalent en Joules de la calorie ($W = JQ$ avec $J = 4,1868\,\mathrm{J.Cal}^{-1}$), les unités "électrostatiques" aux unités "magnétiques" en fixant celle de μ_0 et le mètre à la seconde en fixant celle de c. Ce processus de rattachement de grandeurs entre elles n'a chaque fois été rendu possible que par des découvertes physiques préalables (équivalence travail "chaleur", unification de l'électromagnétisme, relativité) ; il n'est donc pas terminé. D'ailleurs, anticipant sur la métrologie, de nombreux physiciens considèrent l'entropie comme une grandeur sans dimension ; ceci revient à considérer la constante de Boltzmann comme sans dimension et à attribuer à la température thermodynamique non plus une dimension propre mais celle d'une énergie (*cf.* $\frac{1}{2}\,m < v^2 >= \frac{3}{2}\,k_B T$ ou le facteur de Boltzmann $\exp(\frac{-E}{k_B T})$).

▷ Dimensions et domaines de la physique

Le nombre de "dimensions de base", donc de degrés de liberté pour les changements d'unités, dépend du domaine de la physique considéré ; ce nombre est d'autant plus petit que l'étendue de la validité de la théorie considérée est grande. Ainsi, il diminue d'une unité lorsque l'on passe de la **physique classique galiléenne** ("vitesses"$\ll c$, "actions"$\gg \hbar$) au domaine plus général de la **physique classique einsteinienne** ("vitesses" quelconques, "actions"$\gg \hbar$). Ceci est dû à l'apparition d'une nouvelle constante de la physique avec la relativité d'Einstein, à savoir la vitesse invariante c de la lumière dans le vide, qui permet d'identifier les grandeurs masse et énergie au travers de la relation $E = mc^2$; en posant $[c] = 1$ (de même qu'on a posé $[J] = 1$ pour identifier travail et "chaleur" et avoir $[W] = [Q]$) on a $[E] = [M]$, $[L] = [T]$, $[\mu_0] = [\epsilon_0]^{-1}\ldots$ L'apparition de la constante $\hbar$ avec la **physique quantique** permet quant à elle d'identifier énergie et pulsation et quantité de mouvement et vecteur d'onde au travers des relations de Planck-Einstein $E = \hbar\omega$ et de de Broglie $\vec{p} = \hbar\vec{k}$; en posant $[\hbar] = 1$ on a $[E] = [T]^{-1}$ et $[P] = [L]^{-1}$. Finalement en **physique quantique relativiste** où interviennent à la fois c et $\hbar$, et considérant par ailleurs que les charges électriques sont sans

dimension car multiples d'une charge fondamentale q_e, il ne reste plus qu'une dimension : on a en effet $[E] = [M] = [T]^{-1} = [L]^{-1}$ si on pose $[\hbar] = [c] = 1$. Les physiciens des particules ont même pris l'habitude d'"oublier" les constantes c et $\hbar$ dans les formules, écrivant par exemple $\lambda_c = T_c = m^{-1}$ pour la longueur d'onde ou temps Compton, ou $E_r = \dfrac{qa}{4\pi\epsilon_0 r}\sin\theta$ pour le champ rayonné par une charge accélérée ; ils mesurent aussi les masses en Gev et les longueurs en Gev^{-1}. Ceci est sans conséquence pour un physicien "ordinaire" qui peut toujours rétablir les expressions complètes incluant c et $\hbar$ par l'analyse dimensionelle. En effet $\lambda = m^{-1}\hbar^\alpha c^\beta$ de dimension $[EL^{-2}T^2]^{-1}[ET]^\alpha[LT^{-1}]^\beta$ n'est une longueur que pour $\alpha = 1$ et $\beta = -1$ et par conséquent $\lambda_c = \hbar(mc)^{-1} = cT_c$; on rétablit de même l'expression $E_r = \dfrac{qa}{4\pi\epsilon_0 rc^2}\sin\theta$ par comparaison à $E_{es} = \dfrac{q}{4\pi\epsilon_0 r^2}$. Par ailleurs si une masse m vaut 1Gev, c'est que $mc^2 = 1,6\times10^{-19}\times10^9$ J, d'où on déduit m en kg sachant que $c = 3\times10^8\,\text{m.s}^{-1}$.

REMARQUE. En **théorie quantique des champs** (relativiste) la connaissance des dimensions des champs, $[E]$ pour les bosons (*cf.* les potentiels électromagnétiques V et $\overrightarrow{A}$) et $[E]^{\frac{3}{2}}$ pour les fermions ($[\Psi]^2\,[L]^3 = 1$), ainsi que l'exigence que les constantes de couplage des interactions fondamentales soient sans dimension, ont joué un rôle prédictif important pour la théorie électrofaible et la chromodynamique. Par contre comme la constante de couplage G de la **gravitation** a une dimension, $[G] = [E]^{-2} = [L]^2$, cette théorie est plutôt considérée comme phénoménologique, approximation d'une théorie fondamentale nécessaire à l'échelle de Planck $l_{\text{Planck}} = G^{\frac{1}{2}}$ $(= (\hbar G c^{-3})^{\frac{1}{2}} = 10^{-35}$ m en rétablissant $\hbar$ et $c)$.

1.3.2 Applications et limites de l'analyse dimensionnelle

▷ Lois "devinées" ; ordres de grandeurs

EXEMPLE CLASSIQUE. Un **pendule simple** est caractérisé par sa longueur l et sa masse m et le champ de pesanteur par g. Ces trois grandeurs étant dimensionnellement indépendantes, la période T du pendule s'écrit $T \propto l^\alpha m^\beta g^\gamma$; comme $[g] = [L][T]^{-2}$, on obtient par identification $\beta = 0$, $\gamma = -\alpha = 1/2$. Par rapport à la formule exacte $T = 2\pi\sqrt{l/g}$ il manque le facteur sans dimension 2π, qu'un calcul dimensionnel ne peut évidemment pas donner. Cependant un temps caractéristique pour le pendule, censé rendre compte de la variation plus ou moins rapide de son élongation $\theta(t)$, doit être de l'ordre d'un quart de période (temps nécessaire pour passer de $|\theta_{min}| = 0$ à $|\theta_{max}|$) et la grandeur homogène à un temps qui convient le mieux est l'inverse de la pulsation $\omega^{-1} = T/2\pi$. (Un argument équivalent consiste à dire qu'une variation caractéristique pour un angle, ωt pour le pendule, n'est pas 2π mais un radian). Si le calcul dimensionnel est fait avec ω le résultat pour le pendule est exact et nous verrons qu'il est meilleur en général dans la mesure où le facteur numérique sans dimension est toujours plus proche de 1 qu'avec T (*cf.* exemples 3 Rayonnements ci-dessous).

EXEMPLES MOINS CLASSIQUES. 1) **Atomes et corps solides**. Avec les trois constantes $e^2 = \dfrac{q_e^2}{4\pi\epsilon_0}$, m_e et $\hbar$ qui interviennent pour décrire le comportement d'un électron en physique atomique non relativiste, on peut construire une longueur, une énergie et une vitesse. On "prédit" ainsi le rayon typique d'un atome ou la longueur

d'une liaison $a \propto \dfrac{\hbar^2}{m_e e^2} \simeq 1$Å, l'énergie de liaison atomique $|E_l| \propto m_e e^4/\hbar^2 \simeq$ qques ev, la vitesse v des électrons dans un atome ou un métal $v \propto e^2/\hbar = \alpha c \simeq c/137$. On peut en déduire certaines propriétés des matériaux solides (constitués d'atomes de masse m liés les uns aux autres et distants de a) : leur incompressibilité est d'ordre $\chi^{-1} = -V\dfrac{\delta P}{\delta V} \propto \dfrac{E_l}{a^3} \simeq 10^{11}$ pascal, la vitesse des ondes (sismiques, sonores...) qui s'y propagent est d'ordre $c_s \propto \sqrt{\dfrac{E_l}{m}} \simeq 1\,\mathrm{km.s^{-1}}$, leur fréquence de coupure est d'ordre $\nu_{max} \propto \dfrac{c_s}{a} \simeq 10^{13}\,\mathrm{Hz}$, *etc.*

2) **Corps noir**. On peut prédire que l'entropie (divisée par k_B) d'un corps noir (système quantique de particules de masse nulle relativistes) est proportionnelle à $U^{\frac{3}{4}}V^{\frac{1}{4}}(\hbar c)^{-\frac{3}{4}}$ en remarquant que $UV^{\frac{1}{3}}(\hbar c)^{-1}$ est sans dimension et que S est une grandeur extensive ($S(kU, kV) = kS(U, V)$). L'énergie du corps noir $U \propto V(k_BT)^4(\hbar c)^{-3}$ en fonction de la température (**loi de Stefan**) ou le nombre de photons $N_\gamma \propto V(k_BT/\hbar c)^3 \propto k_B^{-1}S$ s'obtiennent par un raisonnement analogue. Les photons ont alors pour énergie typique k_BT et pour longueur d'onde $\lambda \propto \hbar c/k_BT$. En remplaçant c par la vitesse du son c_s, ces relations sont valables pour les ondes de vibration d'un solide (**phonons**) à condition que $\lambda \geq \lambda_{min} = a$, ce qui revient à $T \leq \hbar c_s/k_Ba \simeq 100\,K$ ou $N_\gamma \leq Va^{-3}$ (nombre d'atomes).

3) **Rayonnements**. La puissance électromagnétique P_{em} rayonnée par une charge électrique q se déplaçant à vitesse uniforme ωR sur un cercle de rayon R dépend à priori de q, R et ω qui caractérisent la particule et son mouvement et des grandeurs $4\pi\epsilon_0$ et c associées au champ électromagnétique. Ces cinq grandeurs n'étant pas dimensionnellement indépendantes, il n'est pas possible d'obtenir l'expression de P_{em} sans hypothèse ou connaissance supplémentaires. Si on s'intéresse à la partie dipolaire du rayonnement il est raisonnable de penser que q et R n'interviennent dans l'expression de P_{em} qu'au travers du moment dipolaire : le produit qR. Si de plus, faisant appel à ses connaissances, le lecteur averti se souvient qu'une puissance est proportionnelle au carré de l'amplitude des ondes et que cette amplitude est elle-même proportionnelle au moment dipolaire alors P_{em} est du type $P_{em} \propto (qR)^2 \times (4\pi\epsilon_0)^\alpha\omega^\beta c^\gamma$. A l'aide de $[P] = [M][L]^{-1}[V]^3$ ($[V] = [L][T]^{-1}$), on obtient $\alpha = -1$, $\beta = 4$, $\gamma = -3$, donc en particulier la proportionnalité de P_{em} à ω^4. Le coefficient numérique exact $2/3$ est proche de 1 ; il aurait été d'ordre $(2\pi)^4 \simeq 10^3$ si on avait fait le calcul dimensionnel avec la période ! Pour la contribution quadrupolaire (puissance proportionnelle au carré du moment quadrupolaire qR^2) on obtient $\alpha = -1$, $\beta = 6$, $\gamma = -5$. Pour les ondes gravitationnelles, de type quadrupolaire, rayonnées par une masse m, on établit de même $P_{grav} \propto Gm^2R^4\omega^6c^{-5}$; le coefficient numérique exact est alors $32/5$.

EXEMPLES "CURIEUX". En cosmologie le rayon de l'univers varie comme $t^{\frac{2}{3}}$ si on suppose qu'il ne dépend que de t , G et M (masse de l'univers). La dépendance en temps du rayon du nuage issu d'une explosion nucléaire s'écrit $r(t) \propto E^{\frac{1}{5}}\rho^{-\frac{1}{5}}t^{\frac{2}{5}}$ si, en plus de t, seules l'énergie libérée E et la masse volumique de l'air ρ interviennent ; on peut alors déterminer E par simple observation de l'évolution de la taille du nuage. De nombreuses lois dites **lois d'échelle**, ayant des comportements en puissances pour les variables temps, espace, *etc.*, se rencontrent en hydrodynamique, transitions de phases, sismologie, physiologie animale, *etc.*

On remarquera que la prédiction dimensionnelle des lois fait appel à une forte intuition physique, puisqu'elle repose essentiellement sur un "bon" inventaire de tous les paramètres susceptibles d'intervenir dans l'expression de la grandeur que l'on cherche à déterminer. Cet inventaire étant rarement aussi évident que dans le cas du pendule simple, les **arguments dimensionnels** ne prennent toute leur valeur (ne serait-ce que comme moyen mnémotechnique) qu'une fois la loi vérifiée expérimentalement et justifiée, au moins en partie, par un modèle mathématique.

▷ Simplification d'équations différentielles ou aux dérivées partielles

EXEMPLE 1. Dans l'équation d'un **oscillateur de Van der Pol**

$$\ddot{x} - a\dot{x}(1 - bx^2) + cx = 0 \ ,$$

les dimensions des paramètres positifs a, b et c sont : $[a] = [T]^{-1}$, $[b] = [x]^{-2}$ et $[c] = [T]^{-2}$. Le changement $x' = xb^{\frac{1}{2}}$, $t' = tc^{\frac{1}{2}}$ permet de récrire l'équation sous forme **adimensionnée**

$$\frac{\mathrm{d}^2x'}{\mathrm{d}t'^2} - 2\epsilon\frac{\mathrm{d}x'}{\mathrm{d}t'}(1 - x'^2) + x' = 0$$

avec $2\epsilon = ac^{-\frac{1}{2}}$. Elle est connue pour posséder un cycle limite $x'_\epsilon(t')$ (vers lequel tend toute solution). L'équation de départ possède donc aussi un cycle limite qui s'écrit $x(t) = b^{-\frac{1}{2}}\, x'_\epsilon(tc^{\frac{1}{2}})$. La dépendance de ce cycle vis-à-vis de b et vis-à-vis de c ou a (à ϵ fixé) se déduit donc de considérations purement dimensionnelles. Seule sa dépendance vis-à-vis du paramètre sans dimension ϵ est non triviale et nécessite une étude analytique ou numérique plus approfondie (*cf.* section 6.5.2).

EXEMPLE 2. Une équation assez générale de l'**hydrodynamique des fluides incompressibles** dans un référentiel tournant est :

$$\rho\left(\partial_t\,\vec{v} + (\vec{v}.\overrightarrow{\nabla})\vec{v}\right) = -\overrightarrow{\nabla}p + \rho\vec{g} + \eta\Delta\vec{v} - 2\rho\overrightarrow{\Omega}\wedge\vec{v} \ .$$

Considérons un écoulement stationnaire dans une enceinte (un tuyau par exemple) de longueur caractéristique L et correspondant à des valeurs V et P de la vitesse et de la pression du fluide en un certain point (à l'entrée du tuyau par exemple) ; supposons de plus que les conditions aux limites sont telles que la solution $\vec{v}_{L,V,P}(\vec{r})$, $p_{L,V,P}(\vec{r})$ est unique. Alors connaissant les dimensions de toutes les grandeurs impliquées, par exemple pour le coefficient de viscosité $[\eta] = [\rho][L]^2[T]^{-1}$ puisque la dimension du laplacien est $[\Delta] = [L]^{-2}$, on peut affirmer en reprenant l'analyse de l'exemple précédent que la solution va s'écrire $\vec{v}_{L,V,P}(\vec{r}) = V\vec{v'}\Big(\frac{\vec{r}}{L}\cdots\Big)$, $p_{L,V,P}(\vec{r}) = Pp'\Big(\frac{\vec{r}}{L}\cdots\Big)$ les « … » représentant l'ensemble des **paramètres sans dimensions** (nombres de Reynolds $\frac{\rho}{\eta}VL$, d'Euler $\frac{2P}{\rho V^2}$, de Froude $\frac{V^2}{gL}$, de Rossby $\frac{V}{\Omega L}$, *etc.*). Une application pratique importante de la dépendance en $\vec{r}L^{-1}$ de la solution est la possibilité de prévoir certains écoulements à l'aide de maquettes déduites par homothétie de la structure réelle ; ces maquettes comme les fluides doivent être choisis pour préserver les valeurs des nombres sans dimension.

▷ Dimensions et ensembles fractaux

Mesurée à partir d'une carte au 1/50000 la longueur de la côte de Bretagne sera plus petite que si on la mesure à partir d'une carte au 1/10000 car, plus l'échelle est grande, plus les détails sont "gommés". Autrement dit si x et x' sont les nombres qui mesurent la longueur d'une courbe avec des unités l et l',on n'aura $x' = x\,\dfrac{l}{l'}$ (c.a.d. $xl = x'l'$), comme on s'y attend "dimensionellement", que pour des courbes suffisamment lisses ; pour la côte de Bretagne la relation sera plutôt de la forme $x' = x\left(\dfrac{l}{l'}\right)^d$ avec $d > 1$. d s'appelle la **dimension fractale** de la côte.

EXEMPLE 1. La courbe de Koch, obtenue comme limite du processus itératif de construction dont les étapes successives sont représentées sur la figure 9, est telle que $x' = 4x$ si $l' = \frac{l}{3}$ d'où $4 = 3^d$ et $d = \frac{\ln 4}{\ln 3}$. De même $d = \frac{\ln 2}{\ln 3}$ pour l'**ensemble de Cantor** (de mesure nulle) obtenu en enlevant le tiers central de $[0, 1]$, puis les tiers centraux des deux segments restants et ainsi de suite.

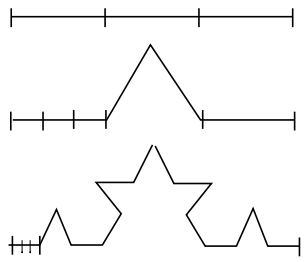


FIGURE 9

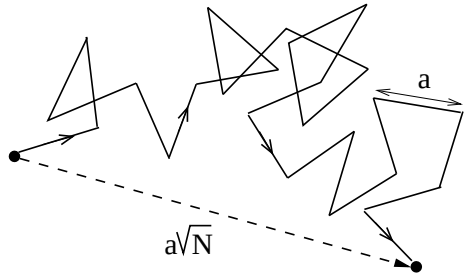


FIGURE 10

EXEMPLE 2. **Trajectoire d'un mouvement brownien**. (figure 10), Sa dimension fractale est toujours $d = 2$ indépendamment du fait que ce mouvement ait lieu sur une ligne, dans un plan ou dans l'espace. On justifie heuristiquement ce résultat en observant que, dans une marche au hasard, on s'éloigne typiquement d'une distance $L = a\sqrt{N}$ du point de départ au bout de N pas de longueur a (*cf.* section 10.3.1). La longueur de la trajectoire correspondant à N pas est donc mesurée par $x = 1$ avec l'unité $l = L$, et par $x' = N = \left(\frac{L}{a}\right)^2$ avec l'unité $l' = a$; par conséquent $x' = x\left(\frac{l}{l'}\right)^2$ et $d = 2$.

EXEMPLE 3. En physique cette notion de dimension fractale intervient pour caractériser les trajectoires asymptotiques dans l'espace de phase de systèmes dynamiques chaotiques (attracteurs étranges) et pour interpréter certaines lois d'échelle qui interviennent dans des phénomènes tels que la percolation, la gélification, la brisure d'une roche, la porosité d'un milieu *etc.* Par exemple, si un milieu matériel est tel que, aux petites échelles, la densité de probabilité de trouver de la matière en $\vec{r}_0 + \vec{r}$ (sachant qu'il y en a en $\vec{r}_0$) obéit à une loi de puissance $g(r) \propto r^{-\alpha}$, on a $d = 3 - \alpha$; en effet, la quantité de matière contenue dans une sphère de rayon R centrée en $\vec{r}_0$, proportionnelle à $\int_0^R r^{-\alpha}\, r^2\, \mathrm{d}r$, varie alors comme $R^{3-\alpha}$, au lieu de R^3 comme habituellement pour un milieu homogène ($\alpha = 0$). Les expériences de diffusion donnent un accès à $g(r)$ (*cf.* section 5.4.2.)

Chapitre 2

Nombres et notation complexes ; plan euclidien

Les nombres complexes $z = x + iy$ ($x, y \in \mathbb{R}$; $i^2 = -1$) ont été introduits dès le $XVI^{\text{ème}}$ siècle sous la forme de nombres "imaginaires" tels que $\sqrt{-1}$ ou $\sqrt[3]{-1}$ pour étendre les formules de résolution des équations algébriques du $2^{\text{ème}}$ et du $3^{\text{ème}}$ degré. A la fin du $XVIII^{\text{ème}}$ siècle Gauss établit le résultat remarquable que tout polynôme $P(z)$ de degré n arbitraire à coefficients complexes admet exactement n racines. Ce résultat, faux si on se limite aux nombres réels (exemple $P(x) = x^2 + 1$), est très important en pratique pour la résolution des Equations Différentielles Linéaires Stationnaires (à coefficients constants) (E.D.L.S.). Une autre vertu de l'ensemble $\mathbb{C}$ des nombres complexes, utile pour la géométrie et la physique à deux dimensions, est la possibilité de coder non seulement les points du plan mais aussi des mouvements (Kepler, harmonique) et des transformations géométriques dans ce plan. Parmi ces transformations les rotations, attachées aux exponentielles complexes $e^{i\theta} = \cos\theta + i\sin\theta$ introduites par Euler, jouent un rôle essentiel dans l'usage de la "notation complexe" en physique classique pour représenter les grandeurs oscillantes (signaux ou ondes polarisées) ainsi que leurs interférences. D'autres applications à la physique linéaire (classique et quantique), et aux E.D.L.S. (étude de stabilité), sont données aux chapitres 4 et 6.

2.1 CALCULS AVEC LES NOMBRES COMPLEXES

2.1.1 Règles de calcul ; exponentielle imaginaire ; fonctions complexes

▷ Rappel des règles de calcul

Soit $z = x + iy$ un nombre complexe. $\overline{z} = x - iy$ est son complexe conjugué, $x = \Re e\, z = \dfrac{z + \overline{z}}{2}$ sa partie reélle, $y = \Im m\, z = \dfrac{z - \overline{z}}{2i}$ sa partie imaginaire et $r = |z| = \sqrt{x^2 + y^2}$ son module. Les opérations addition et multiplication s'effectuent comme pour les réels ;

elles possèdent les propriétés de commutativité, d'associativité et de distributivité, et en plus l'importante propriété de commuter avec l'opération de conjugaison complexe qui consiste à changer i en $-i$. On a notamment

$z_1 + z_2 = x_1 + x_2 + i(y_1 + y_2)$ d'où $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,
$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$ d'où $\overline{z_1 z_2} = \overline{z}_1 \overline{z}_2$,
$z\overline{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2$ d'où $|z_1 z_2| = |z_1||z_2|$,
$|z_1 + z_2|^2 = |z_1|^2 + |z_2|^2 + 2\Re e\,(\overline{z}_1 z_2)$ d'où $|z_1 + z_2| \leq |z_1| + |z_2|$,
$z^{-1} = \dfrac{1}{x + iy} = \dfrac{x - iy}{x^2 + y^2} = \dfrac{\overline{z}}{|z|^2}$ d'où $|z^{-1}| = |z|^{-1}$ et $(\overline{z})^{-1} = \overline{z^{-1}}$.

▷ Exponentielle imaginaire et fonctions sinusoïdales

L'exponentielle imaginaire

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \cdots$$

joue un rôle essentiel dans la théorie et la pratique des nombres complexes en vertu de ses propriétés fondamentales :

$$e^{i(\theta_1+\theta_2)} = e^{i\theta_1} e^{i\theta_2} \quad ; \quad |e^{i\theta}|^2 = e^{i\theta} e^{-i\theta} = 1 \quad ; \quad \left(e^{i\theta}\right)' = i\, e^{i\theta} \ .$$

Elle permet d'introduire algébriquement (sans géométrie) les fonctions **cosinus** et **sinus** par les relations :

$$e^{i\theta} = \cos\,\theta + i\,\sin\,\theta \Longleftrightarrow \cos\,\theta = \Re e\, e^{i\theta} = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin\,\theta = \Im m\, e^{i\theta} = \frac{e^{i\theta} - e^{-i\theta}}{2i} \ .$$

On retrouve à partir des expressions ci-dessus les propriétés classiques de ces fonctions :

$$\cos\,\theta = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n}}{(2n)!} \text{ (fonction paire) ;}$$

$$\sin\,\theta = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!} \text{ (fonction impaire) ;}$$

$\cos\,(\theta_1 + \theta_2) = \cos\,\theta_1\,\cos\,\theta_2 - \sin\,\theta_1\,\sin\,\theta_2$; $\sin\,(\theta_1 + \theta_2) = \sin\,\theta_1\,\cos\,\theta_2 + \sin\,\theta_2\,\cos\,\theta_1$;
$\cos^2\,\theta + \sin^2\,\theta = 1$; $\left(\cos\,\theta\right)' = -\sin\,\theta$; $\left(\sin\,\theta\right)' = \cos\,\theta$;
$\cos^2\,\theta = \left(\dfrac{e^{i\theta} + e^{-i\theta}}{2}\right)^2 = \dfrac{1 + \cos\,2\theta}{2}$; $\cos^3\,\theta = \left(\dfrac{e^{i\theta} + e^{-i\theta}}{2}\right)^3 = \dfrac{\cos\,3\theta + 3\cos\,\theta}{4}$ ... ;
$\cos\,\theta = \sin\,(\theta + \dfrac{\pi}{2}) = -\cos\,(\theta + \pi) = \cos\,(\theta + 2\pi)$;
$\sin\,\theta = \cos\,(\theta - \dfrac{\pi}{2}) = -\sin\,(\theta + \pi) = \sin\,(\theta + 2\pi)$.

Ces dernières propriétés (avance de $\dfrac{\pi}{2}$ de cos θ sur sin θ, périodicité 2π) découlent des relations

$$e^{i\frac{\pi}{2}} = i \quad , \quad e^{i\pi} = i^2 = -1 \ \text{ et } \ e^{2i\pi} = 1 \ ,$$

elles mêmes conséquence de l'étude des variations de cos θ et de sin θ à partir de leurs dérivées ($\dfrac{\pi}{2}$ étant la première valeur positive de θ pour laquelle $\cos\,\theta = 0$).

▷ Argument d'un nombre complexe

L'argument $\theta \in [0, 2\pi[$ (ou phase) de $z = x + iy$ est défini par $x = r\,\cos\,\theta$ et $y = r\,\sin\,\theta$. Il permet d'écrire z sous la forme :

$$z = r(\cos\,\theta + i\,\sin\,\theta) = r\,e^{i\theta} \quad (r = |z|)\ .$$

Cette forme est très utile pour les calculs comme le produit, l'inverse, la recherche des **racines N^{ème}**, nombres complexes $\rho\,e^{i\varphi}$ tels que $\rho^N e^{iN\varphi} = r\,e^{i\theta}$, ou celle des **logarithmes népèriens** (nombres complexes Z tels que $e^Z = r\,e^{i\theta}$) :

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)} \quad ; \quad z^{-1} = r^{-1} e^{-i\theta} \quad ; \quad \text{“}\sqrt[N]{z}\text{”} = r^{\frac{1}{N}} e^{i\frac{\theta}{N}} e^{i\frac{2\pi p}{N}} \ \ p = 0, 1, 2, \ldots, N-1\ ,$$

$$\text{“}\ln\,z\text{”} = \ln\,r + i\theta \text{ modulo } 2\pi i \quad (\ln\,z_1 z_2 = \ln\,z_1 + \ln\,z_2 \text{ modulo } 2\pi i)\ .$$

EXEMPLES : $\sqrt{i} = \pm e^{i\frac{\pi}{4}} = \pm\dfrac{1+i}{\sqrt{2}}$; $\ln\,i = i\,\dfrac{\pi}{2}$ modulo $2\pi i$.

REMARQUE. Les N racines N^{ème} de l'unité $\alpha_p = e^{i2\pi\frac{p}{N}}$ vérifient $\alpha_p^N - 1 = 0$ et donc, en mettant $\alpha_p - 1$ en facteur, $\displaystyle\sum_{n=0}^{N-1} \alpha_p^n = 0$ si $p \neq 0$. Elles servent à définir la **transformée de Fourier discrète** (*cf.* section 5.3.2).

▷ Fonctions de z

En pratique une fonction $f(z)$ est une expression dans laquelle x et y n'apparaissent qu'à travers la combinaison $z = x+iy$, comme par exemple un polynôme $P(z)$, l'exponentielle e^z, la **transformée en z** d'un signal numérique $\{f_n\}$

$$F(z) = \sum_{n=-\infty}^{\infty} f_n\,z^n\ ,$$

mais pas $|z|^2$ dans laquelle apparait aussi $\overline{z} = x - iy$. On fera attention au fait que certaines de ces fonctions comme “$\sqrt[N]{z}$” ou “$\ln\,z$” ne sont pas bien définies même si on choisit une détermination. Par exemple “$\sqrt[N]{z}$” $= r^{\frac{1}{N}} e^{i\frac{\theta}{N}}$ est multiplié par $e^{i\frac{2\pi}{N}}$ quand θ passe continûment de 0 à 2π. Il faut donc toujours préciser le domaine de définition considéré. La **dérivée** au point z de $f(z)$ est définie comme pour les réels par la limite, quand elle existe, de $\dfrac{f(z+\epsilon) - f(z)}{\epsilon}$ lorsque ϵ complexe tend vers zéro (*i.e.* $|\epsilon| \to 0$). (Exemple : $\dfrac{\mathrm{d}z^2}{\mathrm{d}z} = \lim\dfrac{(z+\epsilon)^2 - z^2}{\epsilon} = \lim(2z+\epsilon) = 2z$). Pour toutes les fonctions classiques (définissables localement à partir de séries), les formules de dérivation sont les mêmes que dans le cas des réels : $(e^z)' = e^z$, $(\ln\,z)' = \dfrac{1}{z}$ $(z \neq 0)$, $(\cos\,z)' = -\sin\,z$, *etc.* Des précisions mathématiques et des applications physiques (hydrodynamique, électrostatique...) sont données à la section 8.3.3.

2.1.2 « Théorème fondamental de l'algèbre » et applications

▷ Racines d'un polynôme

Un résultat très important est que tout polynôme $P(z)$ de degré n admet n racines $p_1,\ p_2 \cdots p_n$ et peut donc toujours être factorisé :

$$P(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = a_0 (z-p_1)(z-p_2)\ldots(z-p_n)\ .$$

JUSTIFICATION : si $P(z)$ de degré $n \geq 1$ ne s'annule jamais, $|P(z)|$ atteint une valeur minimale non nulle, disons en p_1. Posant $z = p_1 + \epsilon$ on peut écrire $P(z) = P(p_1) + b_1\epsilon + b_2\epsilon^2 + \cdots + b_n\epsilon^n$. Soit b_k le premier coefficient non nul. Pour $|\epsilon|$ petit on a $|P(z)|^2 \simeq |P(p_1) + b_k\epsilon^k|^2 \simeq |P(p_1)|^2 + 2\Re e\big(\overline{P(p_1)}\, b_k \epsilon^k\big)$, et en faisant "tourner" l'argument de ϵ on peut toujours rendre négative la quantité $\Re e\big(\overline{P(p_1)}\, b_k \epsilon^k\big)$ ce qui contredit l'hypothèse. D'où $P(p_1) = 0$ et la possibilité de factoriser $(z - p_1)$: $P(z) = (z - p_1)Q(z)$. On répète le raisonnement avec $Q(z)$ de degré $n-1$ et ainsi de suite jusqu'à $n = 1$. Remarque : dans le cas réel on ne peut pas jouer sur l'argument de ϵ.

▷ Equations différentielles linéaires stationnaires (E.D.L.S.)

Le résultat ci-dessus s'applique immédiatement à la résolution des équations

$$a_0 f^{(n)}(t) + a_1 f^{(n-1)}(t) + \cdots + a_{n-1} f'(t) + a_n f(t) = 0$$

en permettant de les récrire (puisque $f^{(p)} = \Big(\dfrac{\mathrm{d}}{\mathrm{d}t}\Big)^p f$) :

$$a_0\Big(\frac{\mathrm{d}}{\mathrm{d}t} - p_1\Big)\Big(\frac{\mathrm{d}}{\mathrm{d}t} - p_2\Big)\ldots\Big(\frac{\mathrm{d}}{\mathrm{d}t} - p_n\Big) f(t) = 0\ .$$

Comme les n opérations $\Big(\dfrac{\mathrm{d}}{\mathrm{d}t} - p_j\Big)$ $(j = 1\ldots n)$ commutent entre elles, il est clair sous cette forme que les exponentielles $e^{p_1 t}, \ldots, e^{p_n t}$ sont des solutions particulières de l'équation différentielle. Si toutes les racines p_j sont différentes, les n fonctions exponentielles sont indépendantes et la **solution générale**, qui dépend à priori linéairement de n constantes arbitraires A_j (*cf.* section 6.2.1), est :

$$f(t) = A_1 e^{p_1 t} + A_2 e^{p_2 t} + \cdots + A_n e^{p_n t}\ .$$

Si par contre une racine, par exemple p_1, apparaît k fois il manque $k-1$ solutions. Cependant comme $\Big(\dfrac{\mathrm{d}}{\mathrm{d}t} - p_1\Big)e^{p_1 t} g(t) = e^{p_1 t} g'(t)$ et donc $\Big(\dfrac{\mathrm{d}}{\mathrm{d}t} - p_1\Big)^k e^{p_1 t} g(t) = e^{p_1 t} g^{(k)}(t)$, on voit que à une racine p_1 d'ordre k sont associées les k solutions indépendantes $e^{p_1 t},\ te^{p_1 t},\ t^2 e^{p_1 t}, \ldots, t^{k-1} e^{p_1 t}$.

EXEMPLE : **E.D.L.S. du second ordre** $af'' + bf' + cf = 0$ (a, b et c réels). Soient $p_{1,2} = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ les racines du polynôme " caractéristique" $ap^2 + bp + c$.
1er cas : $b^2 - 4ac > 0$. p_1 et p_2 sont réels et $f(t) = A_1 e^{p_1 t} + A_2 e^{p_2 t}$.
2ème cas : $b^2 - 4ac < 0$. $p_1 = \overline{p_2} = -\lambda + i\omega$ (avec $\lambda = \dfrac{b}{2a}$, $\omega = \dfrac{\sqrt{4ac - b^2}}{2|a|}$) et $f(t) = e^{-\lambda t}\big(A_1 e^{i\omega t} + A_2 e^{-i\omega t}\big)$. Cependant $f(t)$ étant en physique souvent réelle, on préfère écrire la solution sous l'une des deux formes équivalentes (A, B, C réels)

$$f(t) = e^{-\lambda t}\big(A\cos\omega t + B\sin\omega t\big) \quad \text{ou} \quad f(t) = e^{-\lambda t} C\cos(\omega t + \varphi)\ .$$

3ème cas : $b^2 - 4ac = 0$. $p_1 = p_2 = -\lambda$ est une racine double et $f(t) = e^{-\lambda t}(At + B)$.

2.2 PLAN COMPLEXE ET TRANSFORMATIONS ASSOCIÉES

2.2.1 Plan complexe et plan cartésien ; produit scalaire ; aire

Au nombre complexe $z = x + iy = re^{i\theta}$ on associe dans le plan cartésien (xOy) (appelé aussi plan euclidien) un vecteur $\vec{V} = x\hat{x} + y\hat{y}$ de composantes x et y ($\hat{x}$ et $\hat{y}$ étant les vecteurs unitaires des axes) et de norme $|\vec{V}| = |z| = \sqrt{x^2 + y^2}$. De façon équivalente on peut associer à z un point M, dit **affixe** de z, de coordonnées cartésiennes x et y et donc tel que $\overrightarrow{OM} = \vec{V}$ (figure 1). L'addition des nombres complexes $z_1 + z_2$ correspond alors à l'addition vectorielle $\vec{V}_1 + \vec{V}_2$; on a $|\Sigma_i z_i| \leq \Sigma_i |z_i|$ l'égalité n'ayant lieu que si les arguments des z_i sont égaux (figure 2).

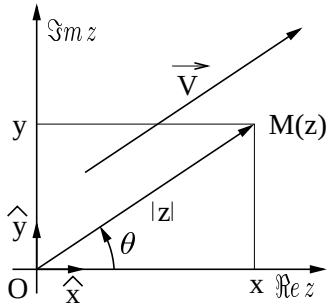


FIGURE 1

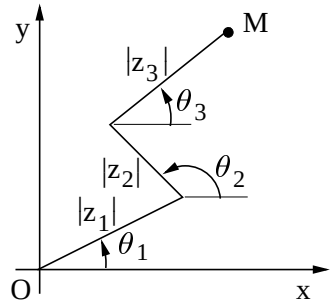


FIGURE 2

Le **produit scalaire** $\vec{V}_1 \cdot \vec{V}_2$ de deux vecteurs et l'**aire algébrique** $\sigma(\vec{V}_1, \vec{V}_2)$ du parallélogramme construit sur $\vec{V}_1$ et $\vec{V}_2$ (figure 3)

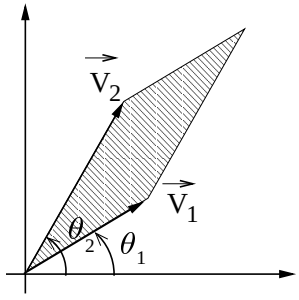


FIGURE 3

sont les quantités invariantes par rotation définies par

$$\vec{V}_1 \cdot \vec{V}_2 = x_1x_2 + y_1y_2 = \Re e(\overline{z}_1 z_2) = |\vec{V}_1||\vec{V}_2| \cos(\theta_2 - \theta_1) = \vec{V}_2 \cdot \vec{V}_1 \ ,$$

$$\sigma(\vec{V}_1, \vec{V}_2) = x_1y_2 - x_2y_1 = \Im m(\overline{z}_1 z_2) = |\vec{V}_1||\vec{V}_2| \sin(\theta_2 - \theta_1) = -\sigma(\vec{V}_2, \vec{V}_1) \ .$$

REMARQUES. L'invariance par rotation est celle relative au changement $\theta_i \to \theta_i + \varphi$. σ est bien l'aire car c'est la seule quantité linéaire en $\vec{V}_1$ et $\vec{V}_2$ qui est nulle si $\vec{V}_1$ et $\vec{V}_2$ sont colinéaires ; (elle est prise égale à 1 pour les vecteurs de base).

EXEMPLES DE PRODUITS SCALAIRES. L'**équation d'une droite** perpendiculaire au vecteur unitaire $\hat{u}$ de composantes $\alpha = \cos\theta_0$ et $\beta = \sin\theta_0$ et distante de d de l'origine s'écrit (figure 4) :

$$\hat{u} \cdot \overrightarrow{OM} = \alpha x + \beta y = d = r\,\cos(\theta - \theta_0) = \Re e\left(e^{-i\theta_0} z\right) .$$

L'égalité $|\overrightarrow{BC}|^2 = |\overrightarrow{BA} + \overrightarrow{AC}|^2$ conduit à la relation bien connue entre les côtés d'un triangle $a^2 = b^2 + c^2 - 2bc\,\cos\hat{A}$.

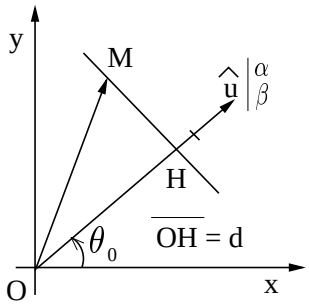


FIGURE 4

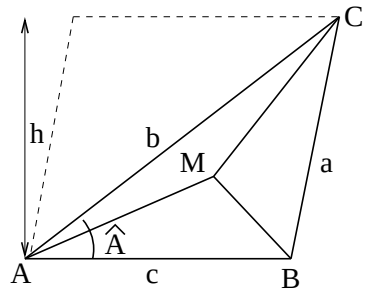


FIGURE 5

EXEMPLES D'AIRES. L'**aire d'un triangle** ABC, étant égale à la moitié de celle d'un parallélogramme construit sur deux de ses côtés (figure 5), vaut $S = \dfrac{1}{2}\,bc\,\sin\widehat{A} = \dfrac{1}{2}\,ab\,\sin\widehat{C} = \dfrac{1}{2}\,ac\,\sin\widehat{B} = \dfrac{1}{2}\,hc$.

Les aires algébriques des triangles orientés MBC, MCA et MAB sont proportionnelles aux **coordonnées barycentriques** λ_A, λ_B et λ_C de M qui vérifient $\lambda_A\,\overrightarrow{MA} + \lambda_B\,\overrightarrow{MB} + \lambda_C\,\overrightarrow{MC} = 0$; par exemple $\sigma\big(\lambda_A\,\overrightarrow{MA} + \lambda_B\,\overrightarrow{MB} + \lambda_C\,\overrightarrow{MC}\,,\,\overrightarrow{MB}\big) = 0$ entraîne $\lambda_A\,\sigma\big(\overrightarrow{MA}\,,\,\overrightarrow{MB}\big) = \lambda_C\,\sigma\big(\overrightarrow{MB}\,,\,\overrightarrow{MC}\big)$ soit $\dfrac{\text{aire } MAB}{\lambda_C} = \dfrac{\text{aire } MBC}{\lambda_A}$. En thermodynamique le "point" triple (phases A, B et C) est représenté par l'intérieur d'un triangle dans le diagramme Volume-Entropie (ou Enthalpie) et M correspond à un **mélange triphasé** de composition $(\lambda_A, \lambda_B, \lambda_C)$.

En cinématique, la **vitesse aréolaire** $\dot{S}(t)$ qui est l'aire balayée par unité de temps par un vecteur $\overrightarrow{0M}(t) = \vec{r}(t)$ s'écrit (figure 6) :

$$\dot{S} = \frac{1}{2}\,\sigma\big(\vec{r}, \frac{\mathrm{d}\vec{r}}{\mathrm{d}t}\big) = \frac{1}{2}\,\Im m\,(\overline{z}\dot{z}) = \frac{1}{2}\,r^2\dot{\theta} .$$

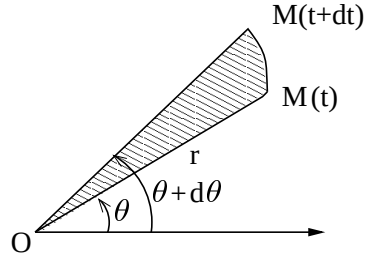


FIGURE 6

2.2.2 Transformations dans le plan complexe

▷ **Transformations euclidiennes** (figure 7)

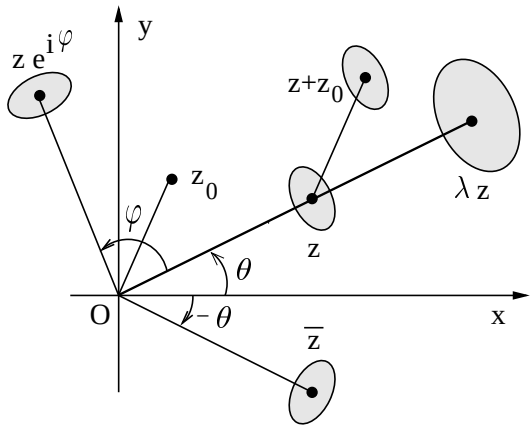


FIGURE 7

La simple observation de la figure montre que l'addition $z \to z + z_0$ correspond à une **translation** de l'affixe M de z, la multiplication par un réel $z \to \lambda z$ à une **homothétie** de centre O et de rapport λ (**symétrie ponctuelle** ou rotation de π si $\lambda = -1$), la multiplication par une exponentielle imaginaire $z \to e^{i\varphi} z$ à une **rotation** de centre O et d'angle φ. Leur composition $z \to az + b$ est une **similitude**. Enfin l'opération $z \to \overline{z}$ est une **symétrie par rapport à Ox**.

▷ **Inversion** (figure 8)

En géométrie l'inversion de centre O et de rapport λ (réel positif) correspond à $r \to \lambda/r$ et $\theta \to \theta$, donc à $z \to \lambda/\overline{z}$. Nous appellerons " inversion" (entre guillemets) l'opération $z \to \lambda/z$ donc $r \to \lambda/r$ et $\theta \to -\theta$ qui en diffère par une symétrie d'axe Ox.

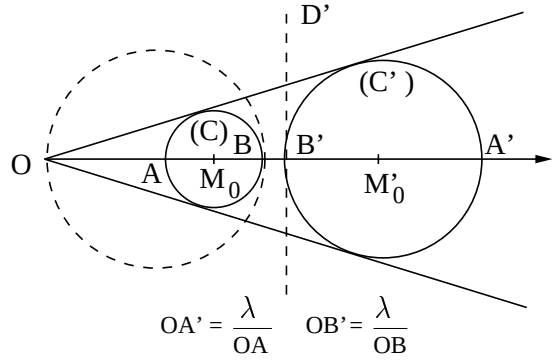


FIGURE 8

Propriété principale : l'inversion de centre O transforme les cercles en "droites ou cercles" et les droites en "droites ou cercles" selon que ces courbes passent ou non par O.

DÉMONSTRATION : l'équation d'un cercle de rayon R et de centre M_0 est $(x-x_0)^2+(y-y_0)^2 = |z-z_0|^2 = R^2$ qui s'écrit $z\overline{z}-\overline{z}_0 z-z_0\overline{z}+|z_0|^2-R^2=0$, ou encore plus généralement $a\,z\overline{z}+\overline{\alpha}z+\alpha\overline{z}+b=0$ (avec a et b réels). Sous cette dernière forme, qui inclut le cas limite des droites lorsque $a=0$, on voit que l'inversion $z \to \frac{1}{\overline{z}}$ conduit à l'équation $bz\overline{z}+\overline{\alpha}\,z+\alpha\overline{z}+a=0$, qui est d'un cercle si $b\neq 0$ ou d'une droite si $b=0$; les centres M_0 $\left(z_0 = -\frac{\alpha}{a}\right)$ et M_0' $\left(z_0' = -\frac{\alpha}{b}\right)$ sont alignés avec O. On remarquera sur la figure 8 que, quand A tend vers O (B restant fixe), le cercle C' inverse du cercle C tend vers la droite D'.

La composition des similitudes et de l'"inversion" conduit aux **homographies** $z \to \dfrac{az+b}{cz+d}$. On vérifie que leur loi de composition

$$z \to z_1 = \frac{a_1 z + b_1}{c_1 z + d_1} \to z_2 = \frac{a_2 z_1 + b_2}{c_2 z_1 + d_2} = \frac{az+b}{cz+d}$$

équivalente à la multiplication des matrices 2×2 :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a_2a_1+b_2c_1 & a_2b_1+b_2d_1 \\ c_2a_1+d_2c_1 & c_2b_1+d_2d_1 \end{pmatrix} .$$

Les homographies préservent aussi le caractère "droite ou cercle".

▷ Transformations conformes

Ce sont les transformations $z \to z' = f(z)$ associées aux fonctions complexes $f(z)$ dérivables. Elles ont la propriété remarquable de conserver les angles ; en particulier tout **réseau de courbes orthogonales** est transformé en un réseau de même nature.

DÉMONSTRATION (figure 9) : soient deux courbes C_1 et C_2 se coupant en M (affixe de z) et M_i $(i=1,2)$ deux points voisins de M, affixes de $z_i = z+\epsilon_i$, situés sur C_i. Leurs transformés M_i' sur C_i' sont les affixes de $z_i' = f(z_i) \simeq f(z)+\epsilon_i\, f'(z)$. L'angle $(\overrightarrow{MM_1}, \overrightarrow{MM_2})$ entre C_1 et C_2 est $\theta \simeq \arg\,\epsilon_2 - \arg\,\epsilon_1$ et celui entre C_1' et C_2' est $\theta' \simeq \arg\, f'(z)\epsilon_2 - \arg\, f'(z)\epsilon_1$; comme $\arg\, f'(z)\epsilon_i = \arg\, f'(z) + \arg\,\epsilon_i$, on a bien $\theta = \theta'$.

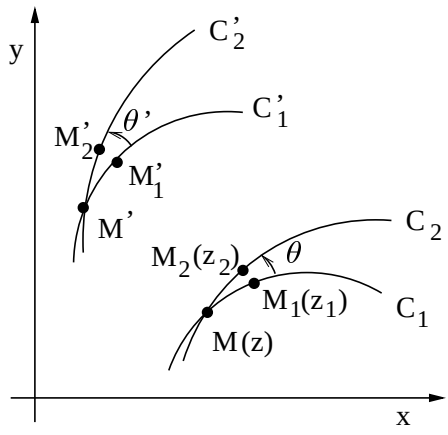


FIGURE 9

Application. Les lieux géométriques de M définis par

$$\text{angle}\,(\overrightarrow{MB},\overrightarrow{MA}) = \text{ constante (modulo } \pi) \quad \text{ et } \quad \frac{MA}{MB} = \text{ constante}$$

sont deux **réseaux de cercles**, les premiers passant par A et B et les seconds orthogonaux aux premiers (figure 12).

DÉMONSTRATION : l'homographie $z \to z' = \frac{z-z_A}{z-z_B}$ qui amène A en O et B à l'infini transforme les lieux cherchés en respectivement les droites passant par O ($\arg z' =$ constante modulo π) et les cercles centrés en O ($|z'| =$ constante).

EXEMPLE 1. **Optique et points de Weierstrass**. Pour un dioptre, de point courant M, deux points A et A', respectivement dans des milieux d'indice n et 1, sont rigoureusement conjugués si $n\,AM \mp A'M =$ constante (principe de Fermat). Si on veut un dioptre sphérique il faut $n\,AM - A'M = 0$; les points A (réel) et A' (virtuel) sont appelés points de Weierstrass du dioptre. La figure 10 représente une lentille boule d'indice n pour laquelle A et A' sont rigoureusement conjugués.

EXEMPLE 2. **Abaque de Smith**. Un autre exemple d'utilisation de l'"inversion" concerne les ondes sur une ligne ou un guide d'onde. La relation entre le coefficient de réflexion complexe r et l'impédance complexe $Z = X + iY$ placée en bout de ligne (normalisée à l'impédance caractéristique) est $r = \dfrac{1-Z}{1+Z}$ ou $r+1 = \dfrac{2}{1+Z}$. Cette "inversion" de centre $A\,(-1,0)$ transforme les droites $X = 0$ et $X = \mathrm{C}^{\mathrm{ste}} > 0$ dans le cercle de diamètre AB ($|r| = 1$) et dans les cercles intérieurs et tangents en A au précédent ; elle transforme les demi-droites $Y = \mathrm{C}^{\mathrm{ste}}$ ($X > 0$) dans les portions de cercles tangents en A à AB. Ces deux réseaux de cercles orthogonaux (abaque de Smith) permettent de déterminer graphiquement Z connaissant r (figure 11).

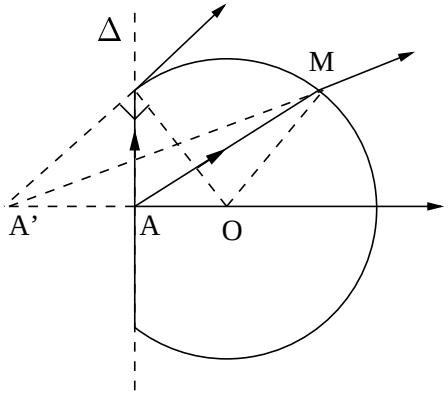


FIGURE 10

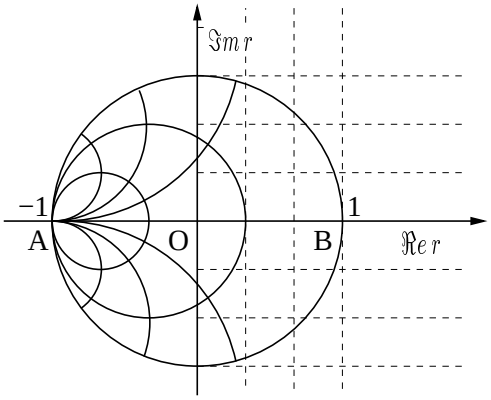
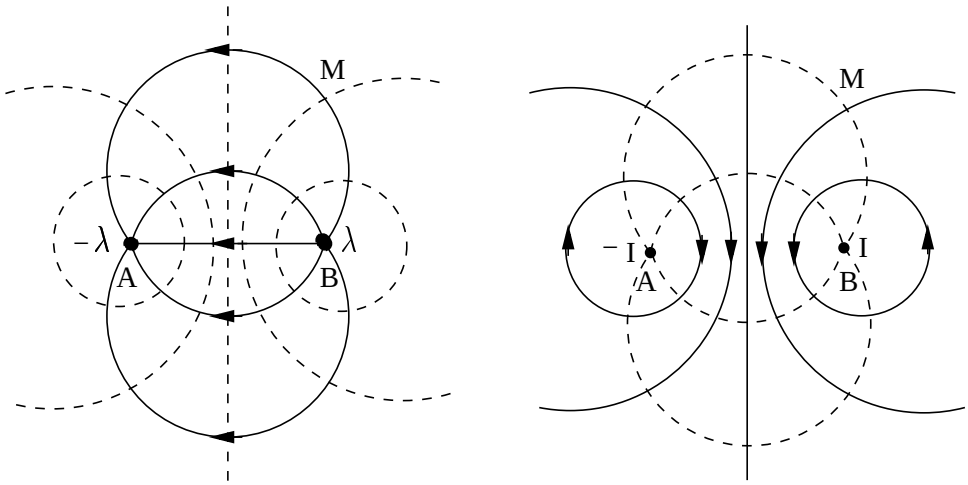


FIGURE 11

EXEMPLE 3. **Electrostatique et magnétostatique de deux fils infinis parallèles** (figures 12a,b). Si ces fils qui percent le plan de la figure en A et B sont chargés (densités $-\lambda$ et λ, figure 12a), ou sont parcourus par des courants (intensités $-I$ et I, figure 12b), les potentiels dont dérivent les champs électrique et magnétique sont respectivement $V_{es}(M) = \dfrac{\lambda}{2\pi\epsilon_0} \ln \dfrac{MA}{MB}$ et $V_{ms}(M) = \dfrac{\mu_0 I}{2\pi} \times \text{angle}(\overrightarrow{MB}, \overrightarrow{MA})$. Les équipotentielles et les lignes de champ sont représentées respectivement en pointillé et en traits pleins sur les figures.



(a) (b)

FIGURE 12

EXEMPLE 4. **Images électriques** (*cf.* section 8.3.2).

2.3 ETUDE DE COURBES ET DE MOUVEMENTS PLANS

2.3.1 Mouvements et courbes en coordonnées polaires

▷ Vitesse et accélération (figure 13)

Le mouvement plan d'un point M généralement décrit par les coordonnées cartésiennes $x(t)$ et $y(t)$ peut aussi l'être par les coordonnées polaires $r(t)$ et $\theta(t)$. Ceci correspond aux deux représentations du nombre complexe z dont M est l'affixe : $z(t) = x(t) + iy(t) = r(t)\, e^{i\theta(t)}$. La vitesse $\vec{v}$ du mouvement, de composantes cartésiennes $\dot{x}$ et $\dot{y}$, et l'accélération $\vec{a}$ de composantes $\ddot{x}$ et $\ddot{y}$ sont les vecteurs associés à $\dot{z}(t)$ et à $\ddot{z}(t)$. En dérivant $r(t)\, e^{i\theta(t)}$ on obtient :

$$\dot{z} = (\dot{r} + ir\dot{\theta})\, e^{i\theta} \quad ; \quad \ddot{z} = \big((\ddot{r} - r\dot{\theta}^2) + i(r\ddot{\theta} + 2\dot{r}\dot{\theta})\big)\, e^{i\theta} \ .$$

Ces expressions peuvent être traduites vectoriellement en introduisant les **vecteurs unitaires radial** $\hat{r} = \dfrac{\vec{r}}{|\vec{r}|}$ associé à $e^{i\theta}$, et **orthoradial** $\hat{\theta}$ associé à $i\, e^{i\theta}$ (déduit de $\hat{r}$ par une

rotation de $\frac{\pi}{2}$). On retrouve les expressions classiques de $\vec{v}$ et $\vec{a}$ en coordonnées polaires :

$$\vec{v} = \dot{r}\,\hat{r} + r\dot{\theta}\,\hat{\theta} \quad ; \quad \vec{a} = (\ddot{r} - r\dot{\theta}^2)\,\hat{r} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\,\hat{\theta}\ .$$

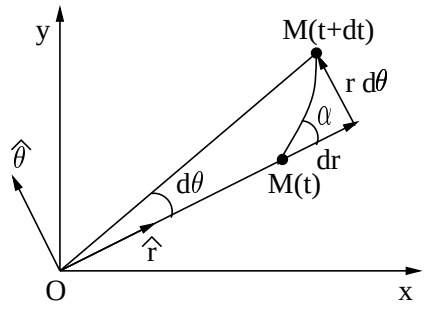


FIGURE 13

▷ Intégrales premières

Pour un mouvement à force centrale où $\vec{a}$ est colinéaire à $\hat{r}$, on a $r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0 = \dfrac{1}{r}\dfrac{\mathrm{d}}{\mathrm{d}t}\,r^2\dot{\theta}$, soit $r^2\dot{\theta} = C = \text{constante}$: c'est la **loi des aires** qui exprime que la vitesse aréolaire est constante. Si en plus la force ne dépend pas de θ ($m\vec{a} = f(r)\,\hat{r}$), on obtient une deuxième intégrale première **énergie** $E = \frac{1}{2}m\vec{v}^2 + V(r)$ (avec $f = -V'$ et V énergie potentielle). En effet $\dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{1}{2}m\vec{v}^2 + V(r)\Big) = m\vec{v}\,.\,\vec{a} + V'(r)\,\dot{r} = \vec{v}\,.\,(m\vec{a} + V'(r)\hat{r}) = 0$. Les équations du mouvement se ramènent alors à deux équations couplées du premier ordre pour $r(t)$ et $\theta(t)$ (en utilisant $v^2 = \dot{r}^2 + r^2\dot{\theta}^2$) :

$$\dot{\theta} = \frac{C}{r^2} \quad ; \quad \frac{1}{2}\,m\dot{r}^2 + V(r) + \frac{mC^2}{2r^2} = E\ .$$

▷ Courbes en coordonnées polaires

Une courbe est définie par une relation $r(\theta)$; par exemple $r = \dfrac{d}{\cos(\theta - \theta_0)}$ pour une droite (figure 4). La tangente à la courbe s'obtient à partir de $\mathrm{d}\vec{r} = \mathrm{d}r\,\hat{r} + r\,\mathrm{d}\theta\,\hat{\theta}$, et l'angle α qu'elle fait avec $\hat{r}$ vérifie $\mathrm{tg}\,\alpha = r\dfrac{\mathrm{d}\theta}{\mathrm{d}r}$ (figure 13).

2.3.2 Coniques en coordonnées polaires et cartésiennes ; foyers

Les coniques apparaissent en géométrie et en physique sous trois formes principales.

▷ Coniques en coordonnées polaires (figure 14)

Elles sont définies par

$$r(\theta) = \frac{p}{1 + e\,\cos\theta} \quad (p > 0\ ;\ e \geq 0)\ ,$$

les angles étant mesurés à partir de l'axe de la conique $\theta = 0$.

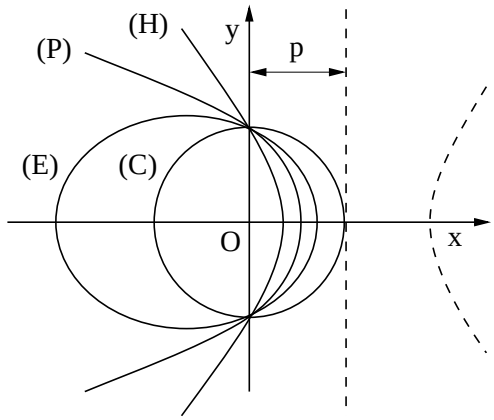


FIGURE 14

Les **cercles** (C), **ellipses** (E), **paraboles** (P) et **hyperboles** (H) correspondent respectivement aux valeurs $e = 0$, $e \in]0,1[$, $e = 1$ et $e > 1$ de l'**excentricité** e (sur la figure $e = 1/2$ pour (E) et $e = 3/2$ pour (H)). Pour p fixé et $r > 0$ (voir la remarque ci-dessous), toutes ces courbes passent par les points $r = p$, $\theta = \pm\dfrac{\pi}{2}$. Elles sont fermées si $e < 1$; si $e \geq 1$ elles présentent des branches infinies dans les directions $\pm\theta_\infty$ telles que $\cos\theta_\infty = -\dfrac{1}{e}$.

▷ Coniques en coordonnées cartésiennes

Leur équation $x^2 + y^2 = (p - ex)^2$ s'obtient à partir de $r = p - e\,r\cos\theta$. Pour $e = 1$ on obtient l'équation d'une parabole $x = \dfrac{p}{2} - \dfrac{y^2}{2p}$. Pour $e \neq 1$ on regroupe les termes en x et x^2 dans un carré : $(1 - e^2)\Big(x + \dfrac{ep}{1 - e^2}\Big)^2 + y^2 = \dfrac{p^2}{1 - e^2}$. Cette expression montre l'existence d'un **centre de symétrie** C de coordonnées $y_{_C} = 0$ et $x_{_C} = -\dfrac{ep}{1 - e^2}$; $x_{_C} = 0$ pour les cercles, $x_{_C} = -c < 0$ pour les ellipses (figure 15) et $x_{_C} = c' > 0$ pour les hyperboles (figure 16 avec $\overline{OC} = c'$).

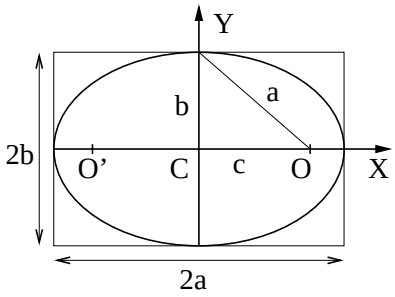


FIGURE 15

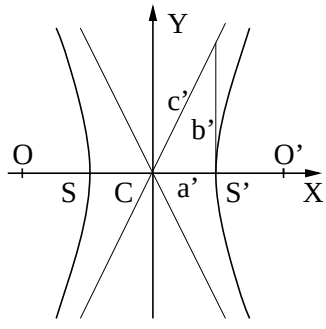


FIGURE 16

Avec C comme origine des coordonnées, on a

$$\text{pour les ellipses :} \quad \frac{X^2}{a^2}+\frac{Y^2}{b^2}=1 \quad ; \quad a=\frac{p}{1-e^2} \quad ; \quad b=a\sqrt{1-e^2} \quad ; \quad e=\frac{c}{a}<1\,,$$

$$\text{pour les hyperboles :} \quad \frac{X^2}{a'^2}-\frac{Y^2}{b'^2}=1 \quad ; \quad a'=\frac{p}{e^2-1} \quad ; \quad b'=a'\sqrt{e^2-1} \quad ; \quad e=\frac{c'}{a'}>1\,.$$

Les interprétations géométriques des coefficients a, b, c, a', b', c' et des relations $a^2=b^2+c^2$ et $c'^2=a'^2+b'^2$ sont illustrées sur les figures.

L'ellipse se déduit du cercle de rayon a par une affinité de rapport $\dfrac{b}{a}$ suivant Oy, d'où son aire $\pi a^2\,\dfrac{b}{a}=\pi ab$. L'hyperbole a pour asymptotes $Y=\pm\dfrac{b'}{a'}\,X$.

REMARQUE. Lorsque $e>1$ la courbe d'équation $r=p(1+e\cos\theta)^{-1}$ avec $r>0$ ne correspond qu'à la branche d'hyperbole "la plus proche" de O ; l'autre branche (en pointillé sur la figure 14) correspond à l'extension de la formule à des valeurs négatives de r.

Paramétrisation avec origine au centre. Elle s'écrit (*cf.* $\cos^2\varphi+\sin^2\varphi=1$ et $\cosh^2\phi-\sinh^2\phi=1$) :

$$X=a\cos\varphi\,,\ Y=b\sin\varphi\ \text{(ellipses)} \quad ; \quad X=\pm a'\cosh\phi\,,\ Y=b'\sinh\phi\ \text{(hyperboles)}\,.$$

Une propriété importante de cette paramétrisation pour l'ellipse est que l'aire balayée par le vecteur $\overrightarrow{CM}$ est proportionnelle à $\Delta\varphi$; en effet $\mathrm{d}S=\dfrac{1}{2}\,\sigma\,(\overrightarrow{CM},\mathrm{d}\overrightarrow{CM})=\dfrac{1}{2}(X\,\mathrm{d}Y-Y\,\mathrm{d}X)=\dfrac{ab}{2}\,\mathrm{d}\varphi$ (figure 17 ; l'aire hachurée vaut $\dfrac{1}{2}\,ab\varphi$).

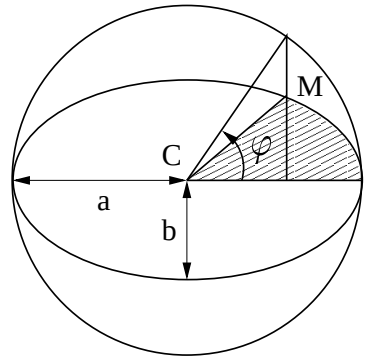
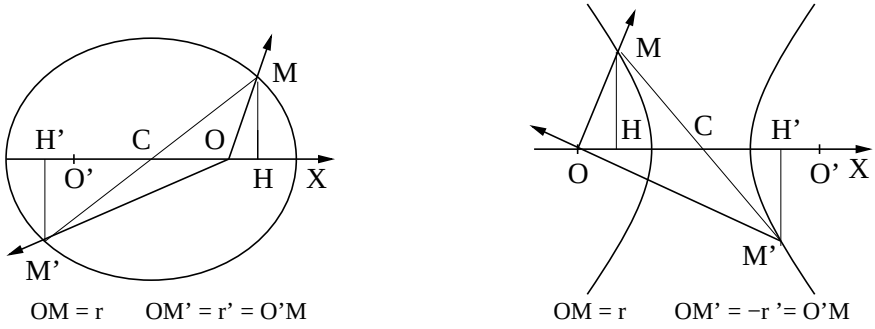


FIGURE 17

▷ Coniques rapportées à leur(s) foyer(s)

Le point O, origine des coordonnées polaires, est appelé **foyer** de la conique. Comme pour les ellipses et les hyperboles le point O' symétrique de O par rapport à C joue un rôle identique, ces coniques possèdent deux foyers O et O'. Considérons alors deux points

M et M' de la conique symétriques par rapport à C et leurs projections H et H' sur OO'. Pour une ellipse (figure 18a) on a $r + e\,\overline{OH} = r' + e\,\overline{OH'} = p$, soit $r + r' = 2p - 2e\,\overline{OC} =$ constante. Pour une hyperbole (figure 18b) ceci reste vrai mais $r' < 0$ (*cf.* remarque ci-dessus),



FIGURES 18

et donc $r - |r'| =$ constante. Pour déterminer les constantes on prend M sur l'axe OX. On en déduit qu'ellipses et hyperboles sont les lieux géométriques des points M tels que :

$$MO + MO' = 2a \quad (\textbf{ellipses}) \quad ; \quad |MO - MO'| = 2a' \quad (\textbf{hyperboles}) \; .$$

A partir de la relation $r + \overline{OH} = p$, on établit de même qu'une **parabole** est le lieu des points M dont la somme de la distance à O et de la distance mesurée algébriquement à une droite fixe est constante :

$$MO + \overline{MK} \quad \text{et} \quad MO + \overline{MK}\,' \quad \text{sont constants (parabole ; figure 19).}$$

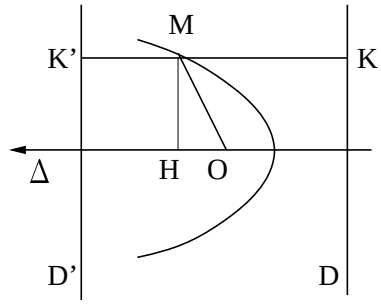
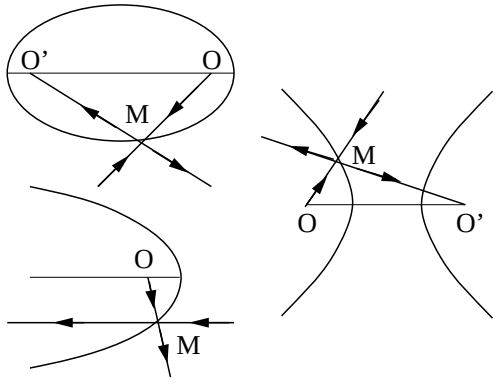


FIGURE 19

Une parabole est une ellipse ou une hyperbole dont un foyer est à l'infini.

Applications à l'optique. En **optique géométrique** le principe de Fermat dit que deux points O et O' sont conjugués par rapport à un miroir (point courant M) si $MO \pm MO' =$ constante (signe moins si les caractères réel ou virtuel de O et O' ne sont pas les mêmes). Conséquence : les surfaces de révolution engendrées par les coniques sont des miroirs parfaits pour leurs foyers (figures 20). En **optique ondulatoire** les



FIGURES 20

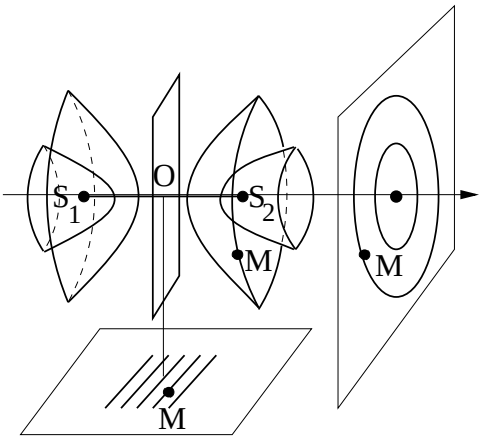


FIGURE 21

interférences de deux sources S_1 et S_2 monochromatiques et cohérentes présentent un déphasage fixe sur les hyperboloïdes de révolution $S_1M - S_2M =$ constante lorsque la lumière arrivant en M vient de S_1 et S_2. Les intersections de ces hyperboloïdes avec des plans d'observation peuvent donner des cercles (observation dans la direction S_1S_2), ou des branches d'hyperboles assimilables à des droites parallèles (observation à grande distance dans une direction perpendiculaire à S_1S_2) (figure 21).

2.3.3 Mouvement de Kepler

Ce mouvement plan (*cf.* section 3.3.2) à force centrale attractive en $\dfrac{1}{r^2}$ est décrit par les équations différentielles, pour le rayon vecteur $\vec{r}$ joignant deux corps ou pour $z = x + iy$,

$$\ddot{\vec{r}} = -\alpha\,\frac{\hat{r}}{r^2} \quad \text{ou} \quad \ddot{z} = -\alpha\,\frac{e^{i\theta}}{|z|^2} \quad (\alpha > 0)\ .$$

($\alpha = G(m_1 + m_2)$ pour le problème gravitationnel ; *cf.* section 3.3.2). Compte tenu de la loi des aires $r^2\,\dot{\theta} = C$, l'équation pour z devient $\ddot{z} = i\,\dfrac{\alpha}{C}\,\dfrac{\mathrm{d}}{\mathrm{d}t}e^{i\theta}$ (ou $\ddot{\vec{r}} = \dfrac{\alpha}{C}\,\dfrac{\mathrm{d}}{\mathrm{d}t}\hat{\theta}$) et s'intègre en $\dot{z} = i\,\dfrac{\alpha}{C}\,e^{i\theta} +$ constante (ou $\dot{\vec{r}} = \dfrac{\alpha}{C}\,\hat{\theta} + \overrightarrow{C^{ste}}$). On peut toujours, en changeant z en $z\,e^{i\theta_0}$ (rotation), se ramener à :

$$\dot{z} = i\,\frac{\alpha}{C}\,(e^{i\theta} + e) \ \ (e \text{ réel positif}) \quad \text{ou} \quad \vec{v} = \frac{\alpha}{C}\,(\hat{\theta} + e\,\hat{y})\ .$$

Cette expression donne la direction de la vitesse en tout point (figure 22), et conduit à l'équation de la **trajectoire** en écrivant $C = \Im m(\overline{z}\dot{z}) = \dfrac{\alpha r}{C}\,\Re e(1 + e\,e^{-i\theta})$ (ou $C = \sigma(\vec{r}, \vec{v})$) :

$$r(\theta) = \frac{p}{1 + e\,\cos\theta} \quad \text{avec} \qquad p = \frac{C^2}{\alpha} > 0\ .$$

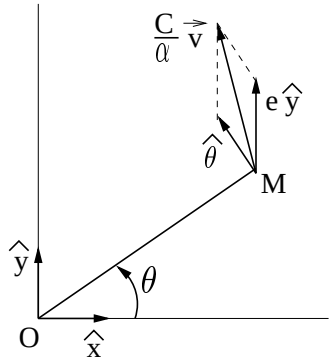


FIGURE 22

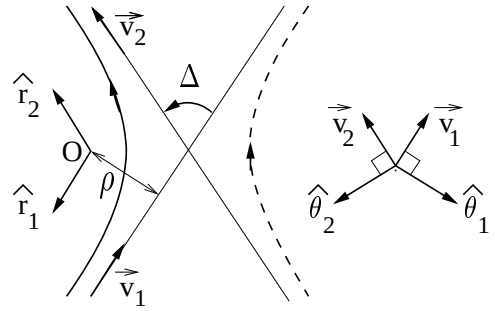


FIGURE 23

On retrouve les équations en polaire des coniques. Dans le cas des ellipses on obtient la **période** T du mouvement en écrivant que l'aire $\frac{1}{2}\,CT$ balayée pendant une période est la surface πab de l'ellipse. Comme $b = a\sqrt{1-e^2}$ et $C^2 = \alpha p = \alpha a(1-e^2)$, on a :

$$T = \frac{2\pi}{\sqrt{\alpha}}\,a^{\frac{3}{2}} \qquad (3^{\text{ème}} \text{ loi de Kepler})\ .$$

Dans le cas des hyperboles (figure 23), l'**angle de diffusion** Δ entre les vitesses initiale $\vec{v}_1$ et finale $\vec{v}_2$ se déduit de l'expression des vitesses en écrivant $\vec{v}_2 - \vec{v}_1 = \frac{\alpha}{C}\,(\hat{\theta}_2 - \hat{\theta}_1)$; des relations $|\vec{v}_1| = |\vec{v}_2| = v$ (conservation de l'énergie), $|\vec{v}_2 - \vec{v}_1| = 2v\,\sin\frac{\Delta}{2}$ et $|\hat{\theta}_2 - \hat{\theta}_1| = 2\,\cos\frac{\Delta}{2}$, et de $C = \rho\,v$ (ρ paramètre d'impact), on déduit :

$$\mathrm{tg}\frac{\Delta}{2} = \frac{\alpha}{\rho v^2}\ .$$

L'énergie E s'obtient à l'aide de $r(\theta)$ et de $v^2(\theta) = \frac{\alpha^2}{C^2}\,(\hat{\theta} + e\hat{y})^2 = \frac{\alpha}{p}\,(1 + e^2 + 2e\,\cos\theta)$:

$$E = \mu\Big(\frac{1}{2}\,\vec{v}^2 - \frac{\alpha}{r}\Big) = \frac{\mu\alpha}{2p}\,(e^2 - 1) \qquad (\mu = \frac{m_1 m_2}{m_1 + m_2} \text{ masse réduite})\ .$$

Pour les ellipses on a $p = a(1-e^2)$ et $E = -\frac{\mu\alpha}{2a} < 0$; pour les hyperboles $E > 0$.

REMARQUES. 1) Lorsque la force est répulsive $\alpha < 0$ (cas de la **diffusion de Rutherford** de particules alpha par un noyau), les calculs précédents sont inchangés mais il faut se rappeler que $r(\theta)$ doit être positif. Comme $p = \frac{C^2}{\alpha} < 0$ ceci n'est possible que si $e > 1$ et la trajectoire est nécessairement une branche d'hyperbole (en pointillé sur la figure 23). 2) L'intégrale première $\dot{z} - i\,\frac{\alpha}{C}\,e^{i\theta}$ correspond, au facteur iC près, au vecteur de Lenz introduit à la section 6.3.

2.3.4 Mouvement harmonique ; vecteurs tournants

L'équation d'un mouvement plan associé à une force centrale attractive proportionnelle à la distance est :

$$\ddot{\vec{r}} = -\omega^2\,\vec{r} \quad \text{ou} \quad \ddot{z} = -\omega^2\,z\ .$$

Sa solution générale qui, pour $z(t)$, s'écrit

$$z(t) = z_+ e^{i\omega t} + z_- e^{-i\omega t} \quad \text{avec} \quad z_\pm = |z_\pm| e^{i\varphi_\pm}\ ,$$

peut être décrite de multiples façons. Elle apparaît ci-dessus comme la somme de deux **vecteurs tournants** (représentés à l'instant $t = 0$ sur la figure 24) ; l'un de longueur $|z_+|$ tourne dans le sens trigonométrique direct (**polarisation circulaire "gauche"** en optique) à la vitesse angulaire ω tandis que l'autre de longueur $|z_-|$ tourne dans le sens inverse (**polarisation circulaire "droite"**) à la vitesse $-\omega$.

En posant $\varphi_\pm = \varphi_0 \pm \varphi$ on a aussi :

$$z(t) = e^{i\varphi_0}\left((|z_+| + |z_-|)\cos(\omega t + \varphi) + i(|z_+| - |z_-|)\sin(\omega t + \varphi)\right)\ .$$

Sous cette deuxième forme, $z(t)$ apparaît comme la somme de deux vecteurs oscillants (**polarisations rectilignes**) qui sont orthogonaux et qui oscillent en quadrature de phase. On en déduit que l'extrémité M de $\overrightarrow{OM}$ (M affixe de z) décrit une ellipse de demi grand axe $OA = a = |z_+| + |z_-|$ incliné de φ_0, de demi petit axe $OB = b = ||z_+| - |z_-||$, et qui est parcourue dans le sens direct si $|z_+| - |z_-| > 0$ (dans le sens inverse si $|z_+| - |z_-| < 0$). Il existe aussi une infinité d'autres façons de décrire $\overrightarrow{OM}(t)$ comme somme de polarisations rectilignes en quadrature mais non orthogonales (figure 25). Il suffit pour cela d'écrire $z(t)$ en fonction des conditions initiales :

$$z(t) = z(t_0)\,\cos\omega(t - t_0) + \dot{z}(t_0)\,\frac{\sin\omega(t - t_0)}{\omega}\ .$$

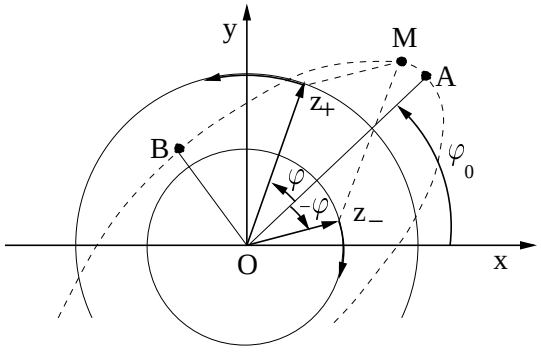


FIGURE 24

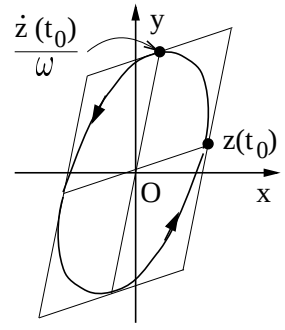


FIGURE 25

REMARQUES. 1) Comme celui de Kepler, le mouvement harmonique est facilement intégrable et possède des trajectoires fermées. En plus des intégrales premières $\frac{1}{2}\,(|\dot{z}|^2 + \omega^2|z|^2)$ (proportionnelle à l'énergie) et $C = \Im m(\overline{z}\dot{z}) = \omega\,(|z_+|^2 - |z_-|^2)$ (dont le signe est bien relié au sens de parcours de l'ellipse), il possède l'intégrale (complexe) $\frac{1}{2}\,(\dot{z}^2 + \omega^2 z^2)$.

2) La représentation d'un vecteur oscillant ($z = z_0\,\cos\omega t$) par deux vecteurs tournants ($\frac{1}{2}\,z_0\,e^{\pm i\omega t}$) est très utilisée en physique (moteur synchrone, RMN, optique...).

2.4 NOTATION COMPLEXE EN PHYSIQUE CLASSIQUE

2.4.1 Signaux réels et complexes

La notation complexe est définie pour les signaux "sinusoïdaux" réels par la correspondance

$$f(t) = a\,\cos(\omega t + \phi) \Longleftrightarrow F(t) = A\,e^{i\omega t} \quad \text{avec } A = a\,e^{i\phi} \text{ amplitude complexe .}$$

Cette correspondance s'étend à tous les signaux réels combinaisons linéaires (sommes ou intégrales) de signaux sinusoïdaux c'est-à-dire, d'après l'analyse de Fourier, à quasiment tous les signaux "utiles" en physique. Nous considérons pour simplifier l'écriture des sommes discrètes :

$$f(t) = a_1\ \cos(\omega_1 t + \phi_1) + a_2\ \cos(\omega_2 t + \phi_2) + \cdots \Longleftrightarrow F(t) = A_1\,e^{i\omega_1 t} + A_2\,e^{i\omega_2 t} + \cdots\ .$$

Le signal complexe $F(t)$ s'appelle aussi **signal analytique** associé à $f(t)$.

REMARQUE. **Conventions et notation complexe**. Le fait d'associer à $e^{i\omega t}$ sa partie réelle $\cos\omega t$ plutôt que sa partie imaginaire $\sin\omega t$ est tout à fait arbitraire. De même on peut faire correspondre à $\cos\omega t$ aussi bien $e^{-i\omega t}$, comme il est fait souvent pour les ondes, que $e^{i\omega t}$, convention utilisée en électronique et en théorie du signal. Une distinction "physique" entre fréquences positives et négatives n'intervient que pour les champs quantiques (*cf.* sections 4.4.4-5).

▷ Propriétés importantes

1) **Linéarité** : $\lambda_1 f_1 + \lambda_2 f_2 + \cdots \Longleftrightarrow \lambda_1 F_1 + \lambda_2 F_2 + \cdots \qquad (\lambda_i \text{ réels})$.
On en déduit en particulier que la somme de signaux sinusoïdaux de même pulsation ω est un signal sinusoïdal de pulsation ω :

$$a_1\,\cos(\omega t+\phi_1)+a_2\,\cos(\omega t+\phi_2)+\cdots = a\,\cos(\omega t+\phi) \quad \text{avec} \quad ae^{i\phi} = a_1e^{i\phi_1}+a_2e^{i\phi_2}+\cdots\ .$$

Ceci se montre en passant aux signaux complexes et en mettant $e^{i\omega t}$ en facteur. Par exemple :

$$a\,\cos\,\omega t + b\,\sin\,\omega t = \Re e\Big((a-ib)e^{i\omega t}\Big) = \sqrt{a^2+b^2}\,\cos(\omega t+\phi) \qquad \text{avec} \quad \phi = \arg(a-ib)\ .$$

(Préciser seulement la valeur de $\operatorname{tg}\phi$, ou de $\sin\phi$ ou de $\cos\phi$ est insuffisant.) On notera que la formule d'addition est encore valable si les amplitudes a_i et les phases ϕ_i dépendent du temps (voir section 2.4.3).

2) **Invariance par translation** : $f(t+\tau) \Longleftrightarrow F(t+\tau)$ d'où $f^{(n)}(t) \Longleftrightarrow F^{(n)}(t)$.
Cette correspondance s'étend aux primitives si on exclut les constantes d'intégration.

3) **Moyenne temporelle** du produit de deux fonctions :

$$< f_1 f_2 > \stackrel{\text{def}}{=} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f_1(t) f_2(t)\,\mathrm{d}t = \frac{1}{2}\,\Re e\big(< \overline{F}_1 F_2 >\big)\ .$$

DÉMONSTRATION : il suffit de le vérifier pour

$$f_1(t) = a_1\,\cos(\omega_1 t + \varphi_1) \ \text{ et } \ f_2(t) = a_2\,\cos(\omega_2 t + \varphi_2)\ ,$$

le cas général correspondant à une "simple extension par linéarité". On a

$$< f_1 f_2 >= \frac{a_1 a_2}{2} \left\langle \cos\Big((\omega_1+\omega_2)t+(\varphi_1+\varphi_2)\Big) + \cos\Big((\omega_2-\omega_1)t+(\varphi_2-\varphi_1)\Big) \right\rangle ,$$

et donc $< f_1 f_2 >= 0$ si $\omega_1 \neq \omega_2$ et $< f_1 f_2 >= \dfrac{a_1 a_2}{2} \cos(\varphi_2-\varphi_1)$ si $\omega_1 = \omega_2$; ce même résultat se retrouve sur l'expression $\dfrac{1}{2}\,\Re e\Big(a_1 a_2 \left\langle e^{i(\omega_2-\omega_1)t}\right\rangle e^{i(\varphi_2-\varphi_1)}\Big)$.

Un exemple important en physique est la **fonction de corrélation** d'un signal $f(t)$:

$$\Gamma_f(\tau) =< f(t) f(t+\tau) >= \frac{1}{2}\,\Re e < \overline{F}(t)\, F(t+\tau) > \ .$$

On vérifie facilement que la fonction de corrélation d'une somme de signaux sinusoïdaux de pulsations différentes est la somme des fonctions de corrélation de chaque signal :

$$f(t) = a_1 \cos(\omega_1 t+\varphi_1)+a_2 \cos(\omega_2 t+\varphi_2)+\cdots \implies \Gamma_f(\tau) = \frac{{a_1}^2}{2} \cos\omega_1\tau+\frac{{a_2}^2}{2} \cos\omega_2\tau+\cdots .$$

2.4.2 Systèmes entrée-sortie ; fonctions de transfert et impédances

La notation complexe est très utilisée dans l'étude des circuits RC, RLC... alimentés en alternatif, et plus généralement en physique lorsqu'on s'intéresse à la réponse d'un système linéaire et stationnaire à une excitation sinusoïdale.

▷ Solutions particulières d'E.D.L.S. avec second membre sinusoïdal

EXEMPLES : $\dot{x}+\lambda x = a\cos\omega t$ et $\ddot{x}+2\lambda\dot{x}+\omega_0^2 x = a\cos\omega t$ $(\lambda \geq 0)$.

On cherche une solution particulière des équations

$$\dot{X}+\lambda X = a\,e^{i\omega t} \ \text{ et } \ \ddot{X}+2\lambda\dot{X}+\omega_0^2 X = a\,e^{i\omega t}$$

proportionnelle à $e^{i\omega t}$. On trouve facilement "de tête" :

$$X_\omega(t) = \frac{a e^{i\omega t}}{i\omega+\lambda} \ \text{ et } \ X_\omega(t) = \frac{a e^{i\omega t}}{(\omega_0^2-\omega^2)+2i\lambda\omega} \ .$$

Les solutions particulières réelles $x_\omega(t)$ s'obtiennent en prenant la partie réelle des $X_\omega(t)$:

$$x_\omega(t) = a\,\frac{\lambda\cos\omega t+\omega\sin\omega t}{\omega^2+\lambda^2} \qquad \text{et} \qquad x_\omega(t) = a\,\frac{(\omega_0^2-\omega^2)\cos\omega t+2\lambda\omega\sin\omega t}{(\omega_0^2-\omega^2)^2+4\lambda^2\omega^2} \ .$$

Remarques. 1) Si $\lambda = 0$ et $\omega = \omega_0$ (résonance sans amortissement), la solution $X_\omega(t)$ n'existe plus. Il faut chercher une solution du type $At\,e^{i\omega_0 t}$. On trouve $X(t) = \dfrac{a}{2i\omega_0}\,t\,e^{i\omega_0 t}$; donc $x(t) = a\,t\,\dfrac{\sin\omega_0 t}{2\omega_0}$ est une solution particulière de $\ddot{x}+\omega_0^2 x = a\cos\omega_0 t$ (avec $x(0) = \dot{x}(0) = 0$). 2) En pratique, pour $\lambda > 0$, ces solutions particulières ("**régime stationnaire**") sont observées lorsque l'excitation $\cos\omega t$ à été appliquée depuis suffisamment longtemps (en toute rigueur à $t = -\infty$) pour que les "**transitoires**" (solutions de l'équation sans second membre) aient disparu.

▷ Systèmes entrée-sortie linéaires et stationnaires

On appelle ainsi une correspondance univoque $e(t) \to s(t)$ entre deux fonctions telle que :

$$\lambda_1 e_1 + \lambda_2 e_2 \to \lambda_1 s_1 + \lambda_2 s_2 \text{ (**linéarité**) et } e(t+\tau) \to s(t+\tau) \text{ (**stationnarité**) .}$$

EXEMPLES. 1) Les équations précédentes récrites $\dot{s} + \lambda s = e$, $\ddot{s} + 2\lambda\dot{s} + \omega_0^2 s = e$ (avec $\lambda > 0$), ou plus généralement toute E.D.L.S. (relative à un système stable) pour $s(t)$, avec $e(t)$ ou ses dérivées au second membre, et avec (attention !) conditions initiales nulles (pour assurer la linéarité), souvent prises à $t = -\infty$. 2) La **moyennisation d'un signal** $e(t)$ ou de façon générale la **convolution** de $e(t)$ avec une fonction $R(t)$ (*cf.* section 5.2.1) :

$$e(t) \longrightarrow s(t) = \frac{1}{\tau}\int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} e(t')\,\mathrm{d}t' \quad ; \quad e(t) \longrightarrow s(t) = \int_{-\infty}^{\infty} R(t-t')\,e(t')\,\mathrm{d}t' \ .$$

Le résultat principal relatif à tous ces exemples est que à toute entrée $e(t)$ sinusoïdale correspond une sortie $s(t)$ sinusoïdale ne pouvant différer de $e(t)$ que par son amplitude et sa phase. En notation complexe et avec l'écriture d'usage en électronique :

$$e(t) = e^{pt} \longrightarrow s(t) = T(p)\,e^{pt} \quad (p = i\omega) \ .$$

JUSTIFICATION (ne faisant appel qu'aux propriétés de linéarité et de stationnarité) : si $e^{i\omega t} \to s_\omega(t)$, on doit avoir $e^{i\omega(t+\tau)} \to s_\omega(t+\tau)$ et $e^{i\omega\tau}e^{i\omega t} \to e^{i\omega\tau}s_\omega(t)$, d'où $s_\omega(t+\tau) = e^{i\omega\tau}s_\omega(t)$ et $s_\omega(\tau) = s_\omega(0)\,e^{i\omega\tau}$. $T(p)$ est la **fonction de transfert** (complexe) du système, appelée aussi **impédance complexe** lorsque $s(t)$ est une tension et $e(t)$ une intensité.

EXEMPLES. Pour les deux E.D.L.S. étudiées ci-dessus, $T(p) = (p+\lambda)^{-1}$ (filtre passe bas du premier ordre) et $T(p) = (p^2 + 2\lambda p + \omega_0^2)^{-1}$ (filtre passe bas du deuxième ordre si $\lambda \gg \omega_0$). Pour la moyennisation on vérifie directement :

$$s_\omega(t) = \frac{1}{\tau}\int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} e^{i\omega t'}\,\mathrm{d}t' = \mathrm{sinc}\,\big(\frac{\omega\tau}{2}\big)\,e^{i\omega t} \quad (\mathrm{sinc}\,x = \frac{\sin x}{x} : \text{ fonction **sinus cardinal**}) ;$$

la moyennisation atténue donc toute fonction sinusoïdale et introduit un déphasage de 0 ou π. Plus généralement pour la convolution, en posant $t'' = t - t'$ et $\omega = 2\pi\nu$, on obtient :

$$s_\omega(t) = \int_{-\infty}^{\infty} R(t-t')\,e^{i2\pi\nu t'}\,\mathrm{d}t' = \hat{R}(\nu)\,e^{i2\pi\nu t} \quad \text{avec} \quad \hat{R}(\nu) = \int_{-\infty}^{\infty} e^{-i2\pi\nu t''}\,R(t'')\,\mathrm{d}t'' \ .$$

La fonction de transfert $\hat{R}(\nu) = T(p = i2\pi\nu)$ est la **transformée de Fourier** de la fonction $R(t)$ (réponse impulsionnelle ; *cf.* section 5.2.3).

REMARQUE. Pour des signaux numériques la relation de convolution entrée sortie est :

$$\{e_n\} \to \{s_n\} \text{ avec } s_n = \sum_m R_{n-m}\,e_m \ .$$

Les transformées en z sont reliées par $\sum_n s_n z^n = S(z) = R(z)\,E(z)$.

2.4.3 Signaux modulés ou quasi-monochromatiques

En physique il est rare que les signaux soient rigoureusement sinusoïdaux c'est-à-dire monochromatiques. Ils se présentent souvent sous la forme d'une somme de signaux monochromatiques ayant des pulsations différentes mais proches d'une valeur ω_0 ; de tels signaux sont dits modulés ou quasi-monochromatiques. Il s'écrivent en notation complexe

$$F(t) = A_1\, e^{i\omega_1 t} + A_2\, e^{i\omega_2 t} + \cdots = \Big(A_1\, e^{i(\omega_1-\omega_0)t} + A_2\, e^{i(\omega_2-\omega_0)t} + \cdots\Big)\, e^{i\omega_0 t} = A_f(t)\, e^{i\omega_0 t}\ ,$$

avec $A_f(t) = a_f(t) + i b_f(t)$ amplitude complexe dépendant du temps, et en notation réelle :

$$f(t) = a_1\, \cos(\omega_1 t + \phi_1) + a_2\, \cos(\omega_2 t + \phi_2) + \cdots = a_f(t)\ \cos\,\omega_0 t - b_f(t)\, \sin\,\omega_0 t\ .$$

Les fonctions $a_f(t)$ et $b_f(t)$ sont lentement variables (comparées à $\cos\,\omega_0 t$).

EXEMPLES. Les **modulation d'amplitude** et **modulation de fréquence** d'un signal "porteur de haute fréquence" $a\,\cos\,\omega_0 t$, par un signal adimensionné $x(t)$ de "basse fréquence" (sensé transmettre l'information), conduisent respectivement aux signaux modulés :

$$a(1 + m x(t))\ \cos\,\omega_0 t \quad \text{et} \quad a\,\cos\big(\omega_0 t + \Delta\omega \int^t x(t')\,\mathrm{d}t'\big)\ ;$$

$m \ll 1$ est le facteur de modulation d'amplitude, et $\omega_0\big(1 + \dfrac{\Delta\omega}{\omega_0}\, x(t)\big)$, avec $\dfrac{\Delta\omega}{\omega_0} \ll 1$, est la pulsation instantanée du signal modulé en fréquence. Les amplitudes complexes sont : $a(1 + m x(t))$ et $a \exp\big(i\Delta\omega \int^t x(t')\,\mathrm{d}t'\big)$.

▷ Fonction de corrélation

Quand on calcule la moyenne temporelle du produit de deux signaux modulés f et g de pulsations proches de ω_0, l'utilisation de la notation complexe montre immédiatement que seules les amplitudes lentement variables interviennent :

$$< f g >= \frac{1}{2}\, \Re e < \overline{F} G >= \frac{1}{2}\, \Re e < \overline{A}_f A_g >= \frac{1}{2}\ < a_f a_g + b_f b_g >\ .$$

En particulier :

$$\Gamma_f(\tau) =< f(t)\, f(t+\tau) >= \frac{1}{2}\, \Re e\big(e^{i\omega_0\tau}\, \Gamma_A(\tau)\big) \quad \text{avec} \quad \Gamma_A(\tau) =< \overline{A}_f(t) A_f(t+\tau) >\ .$$

La fonction de corrélation d'un signal modulé contient une partie rapidement variable (de pulsation ω_0) modulée par la fonction de corrélation lentement variable de l'amplitude complexe.

2.5 APPLICATIONS À L'OPTIQUE ONDULATOIRE

2.5.1 Interférences

Les interférences en optique résultent de l'addition de signaux qui présentent entre eux essentiellement des différences de temps de propagation. La grandeur détectée est une **intensité lumineuse** I qui est proportionnelle à la moyenne temporelle du carré du signal résultant.

▷ Lumière monochromatique

En notation complexe, avec la convention des ondes, un signal monochromatique (période $T = \dfrac{2\pi}{\omega}$, longueur d'onde $\lambda = cT$) s'écrit $A_1 e^{-i\omega t}$. Un autre signal de même pulsation en avance de τ s'écrit $A_2 e^{-i\omega(t+\tau)}$; avance temporelle τ, avance de phase φ et différence de chemin optique δ sont reliés par $\dfrac{\tau}{T} = \dfrac{\varphi}{2\pi} = \dfrac{\delta}{\lambda}$.

L'addition de ces deux signaux correspond à tous les systèmes interférentiels à deux ondes (deux trous, deux fentes, biprisme, *etc.*). Dans ce cas $F(t) = A_1 e^{-i\omega t} + A_2 e^{-i(\omega t+\varphi)}$ et

$$I(\varphi) = \frac{1}{2} < |F(t)|^2 >= \frac{1}{2} |A_1 + A_2 e^{-i\varphi}|^2 = I_0 \left(1 + C \cos(\varphi - \varphi_0)\right) .$$

$I_0 = \dfrac{1}{2}\left(|A_1|^2 + |A_2|^2\right)$ est l'intensité moyenne et $C = \dfrac{2|\overline{A_1} A_2|}{|A_1|^2 + |A_2|^2} = \dfrac{I_{max} - I_{min}}{I_{max} + I_{min}}$ le **facteur de contraste** de la figure d'interférences (figure 26) ; quant à $\varphi_0 = \arg A_2 - \arg A_1$ il rend compte des déphasages supplémentaires liés à des réflexions, au passage par un foyer, *etc.*

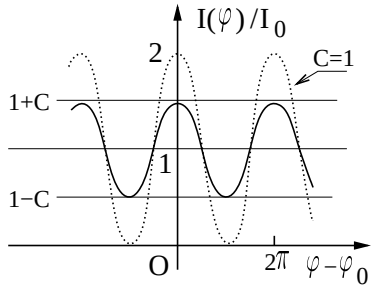


FIGURE 26

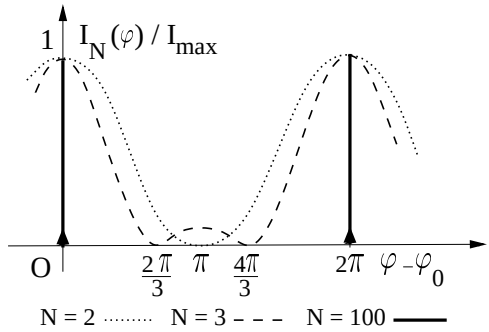


FIGURE 27

L'addition de N signaux également déphasés correspond aux interférences données par des **réseaux**. Dans ce cas $F(t) = A\left(1 + e^{-i\varphi} + \cdots + e^{-i(N-1)\varphi}\right) e^{-i\omega t}$ (série géométrique) et

$$I_N(\varphi) = \frac{1}{2} < |F(t)|^2 >= \frac{1}{2} |A|^2 \left| \frac{1 - e^{-iN\varphi}}{1 - e^{-i\varphi}} \right|^2 = I_{max} \frac{\sin^2 \frac{N\varphi}{2}}{N^2 \sin^2 \frac{\varphi}{2}} ,$$

où $I_{max} = \dfrac{1}{2} N^2 |A|^2$ est obtenue pour $\varphi = 2\pi n$ ($n = 0, 1, 2 \ldots$ est l'**ordre d'interférence**) (figure 27). Pour N grand le rapport $\dfrac{I(\varphi)}{I_{max}}$ est très petit presque partout, sauf pour φ proche de zéro (modulo 2π) où on peut l'approximer par $\operatorname{sinc}^2 \dfrac{N\varphi}{2}$. De part et d'autre du maximum d'ordre n l'intensité s'annule pour la première fois pour $\varphi = 2\pi n \pm \Delta\varphi$ avec $\Delta\varphi = \dfrac{2\pi}{N}$; le rapport $\dfrac{\varphi}{\Delta\varphi} = nN$ qui caractérise la finesse du maximum est le **pouvoir de résolution** du réseau dans l'ordre n (*cf.* cours d'optique).

REMARQUE. Un calcul analogue apparait pour le signal émis par un laser lorsqu'un grand nombre N de modes de la cavité $\omega_n = \omega_0 + n\,\Delta\omega$ sont accrochés en phase. Alors $F(t) = A\sum_{n=0}^{N-1} e^{-i(\omega_0+n\Delta\omega)t}$ et $t\,\Delta\omega$ joue le rôle de φ dans le calcul précédent. Le signal se présente donc comme une succession de pulses intenses espacés dans le temps de $T = \frac{2\pi}{\Delta\omega}$ et de durée $\tau \simeq \frac{T}{N}$.

Le signal à la sortie d'un interféromètre à ondes multiples comme le **Perot Fabry** est une somme discrète infinie de signaux également déphasés et atténués de la forme $F(t) = A\left(1 + Re^{-i\varphi} + R^2e^{-2i\varphi} + \cdots\right)e^{-i\omega t}$; dans cette expression, $R = |R|\,e^{i\varphi_0}$ est un facteur de réflexion de module légèrement inférieur à 1, et le déphasage φ est relié à l'épaisseur e de l'interféromètre et à l'angle d'incidence i (proche de zéro) par $\varphi = 2\pi\,\dfrac{2e\,\cos i}{\lambda}$. L'intensité résultante est

$$I(\varphi) = \frac{1}{2} < |F(t)|^2 >= \frac{1}{2}\,\frac{|A|^2}{|1 - Re^{-i\varphi}|^2} = \frac{I_{max}}{1 + m\,\sin^2\,\frac{\varphi-\varphi_0}{2}}\ ,$$

où $I_{max} = \dfrac{|A|^2}{2(1-|R|)^2}$ est obtenue pour $\varphi - \varphi_0 = 2n\pi$, et où $m = \dfrac{4|R|}{(1-|R|)^2}$ est un coefficient très grand devant 1 ($m \simeq 1600$ pour $|R| = 0,95$) (figure 28). A partir de la relation $I = \dfrac{I_{max}}{2}$ pour $\varphi - \varphi_0 \sim \pm\dfrac{2}{\sqrt{m}}$ (modulo 2π), on déduit la largeur à mi-hauteur des maxima $\Delta\varphi = \dfrac{4}{\sqrt{m}} = 2(1-|R|)$; le **pouvoir de résolution** de l'interféromètre est $\dfrac{\varphi}{\Delta\varphi} = \dfrac{2\pi}{1-|R|}\dfrac{e}{\lambda}$.

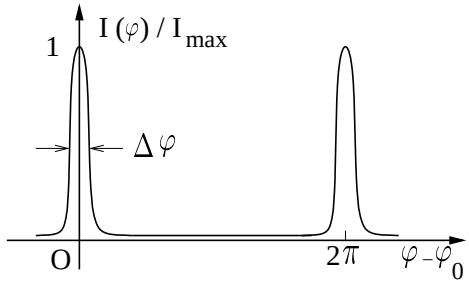


FIGURE 28

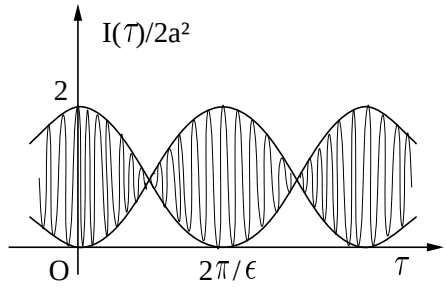


FIGURE 29

▷ Lumière non monochromatique

Dans une interférence à deux ondes $f(t)$ et $f(t+\tau)$, en supposant pour simplifier qu'elles ne diffèrent que par leur temps de propagation, le détecteur mesure l'intensité :

$$I(\tau) =< \big(f(t) + f(t+\tau)\big)^2 >= 2\left(\Gamma_f(0) + \Gamma_f(\tau)\right)\ .$$

Les interférences donnent donc accès à la fonction de corrélation du signal réel et à celle de son amplitude complexe.

EXEMPLE. $f(t) = a\,\cos(\omega_0+\epsilon)t + a\,\cos(\omega_0-\epsilon)t\,;\,\Gamma_f(\tau) = \dfrac{a^2}{2}\cos(\omega_0+\epsilon)\tau + \dfrac{a^2}{2}\cos(\omega_0-\epsilon)\tau$ et $I(\tau) = 2a^2\,(1+\cos\epsilon\tau\ \cos\omega_0\tau)$. Cette expression ressemble au cas monochromatique mais avec un contraste $C(\tau) = \cos\epsilon\tau$ dépendant "lentement" de τ (figure 29). On verra à la section 5.3.3 l'importance de l'étude de $\Gamma_f(\tau)$ pour caractériser les signaux chaotiques.

2.5.2 Diffraction en lumière monochromatique

La **diffraction** correspond aux interférences d'un ensemble "continu" d'ondes et conduit à des calculs d'intégrales.

▷ Diffraction de Fresnel

Si on connait le signal lumineux $A(m)\,e^{-i\omega t}$ en tout point m d'un plan ($z = 0$), le **principe d'Huygens-Fresnel** exprime que le signal $A(M)\,e^{-i\omega t}$ en un point M d'un plan $z > 0$ est une somme de contributions provenant des "points sources" m (figure 30),

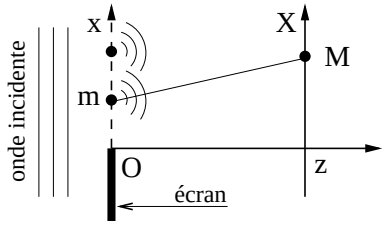


FIGURE 30

chacune contenant un facteur qui rend compte de la propagation du signal de $m(x,y,0)$ à $M(X,Y,z)$. Plus précisément :

$$A(M)\,e^{-i\omega t} = \iint A(m)\,e^{-i\omega t}\,\frac{e^{i2\pi\frac{mM}{\lambda}}}{i\lambda\,mM}\,\mathrm{d}x\,\mathrm{d}y\ ,$$

et dans l'approximation de Gauss ($mM = \sqrt{(X-x)^2+(Y-y)^2+z^2} \simeq z + \dfrac{(X-x)^2+(Y-y)^2}{2z}$)

$$A(M) = \frac{e^{i2\pi\frac{z}{\lambda}}}{i\lambda z}\iint A(x,y)\,e^{i\pi\,\frac{(X-x)^2+(Y-y)^2}{z\lambda}}\,\mathrm{d}x\,\mathrm{d}y\ .$$

EXEMPLE. Pour une ouverture rectangulaire ($x \in [x_1, x_2 = x_1 + l]$ et $y \in [y_1, y_2 = y_1 + L]$) éclairée sous incidence normale ($A(x,y) = A$ constante), l'intégrale factorise et se ramène, par le changement de variables $s = \sqrt{\dfrac{2}{z\lambda}}\,(X-x)$ et $t = \sqrt{\dfrac{2}{z\lambda}}\,(Y-y)$, à :

$$A(M) = A\,\frac{e^{i2\pi\frac{z}{\lambda}}}{2i}\,\big(Z(s_2) - Z(s_1)\big)\big(Z(t_2) - Z(t_1)\big)\ .$$

$Z(s)$ est l'intégrale de Fresnel $Z(s) = \displaystyle\int_0^s e^{i\pi\frac{u^2}{2}}\,\mathrm{d}u$ dont la dépendance en s se lit facilement sur la **spirale de Cornu**, lieu des points affixes de $Z(s)$ (figure 31). On remarquera que $\mathrm{d}s = |\mathrm{d}Z(s)|$ (élément de longueur de la spirale), et que $\dfrac{\mathrm{d}Z}{\mathrm{d}s}$ est alternativement imaginaire et réel pour $s = \sqrt{n}$ $(n = 1, 2, 3...)$. Pour la **diffraction par un demi écran plan** occupant la région $x < 0$ ($t_1 = \infty$ et $t_2 = s_2 = -\infty$), on obtient l'intensité $|A(M)|^2 = \dfrac{|A|^2}{2}\,|Z(-\infty) - Z(\sqrt{\dfrac{2}{z\lambda}}\,X)|^2$ (figure 32).

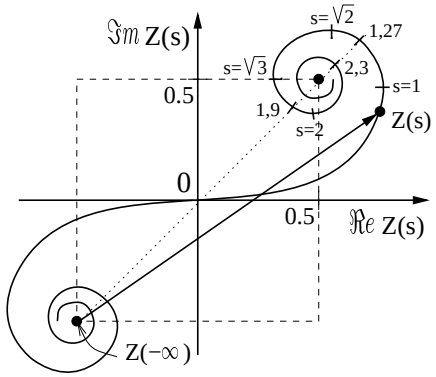


FIGURE 31

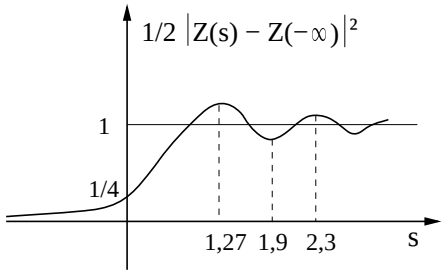


FIGURE 32

▷ **Diffraction de Fraunhofer** (à "l'infini")

Si z est grand devant les dimensions du domaine d'intégration, on peut négliger les termes en $\dfrac{x^2}{z\lambda}$ et $\dfrac{y^2}{z\lambda}$ sous l'intégrale et écrire :

$$|A(M)|^2 \propto \left|\iint A(x,y)\,e^{-2i\pi(ux+vy)}\,\mathrm{d}x\,\mathrm{d}y\right|^2 = |\hat{A}(u,v)|^2 \quad \text{où } u = \frac{X}{z\lambda} \text{ et } v = \frac{Y}{z\lambda}\ .$$

La fonction $\hat{A}(u,v)$ est la **transformée de Fourier** (à deux dimensions) de $A(x,y)$. Pour l'ouverture rectangulaire ci-dessus, et $A(x,y) = A$ constante, on a $|A(M)|^2 \propto |A|^2 \left(\dfrac{lL}{z\lambda}\right)^2 (\operatorname{sinc}\pi ul)^2\,(\operatorname{sinc}\pi vL)^2$ puisque $\left|l^{-1}\int_{x_1}^{x_1+l} e^{-2i\pi ux}\,\mathrm{d}x\right|^2 = (\operatorname{sinc}\pi ul)^2$ (*cf.* section 2.4.2). On verra d'autres exemples à la section 5.4.2.

2.5.3 Polarisations

Lorsque la grandeur physique est vectorielle, la notation complexe définie plus haut se fait composante par composante, et les calculs de moyennes temporelles obéissent aux mêmes règles.

▷ Description des polarisations d'une onde plane électromagnétique monochromatique

Considérons une onde se déplaçant vers les $z > 0$, et notons $\vec{\mathcal{E}}(0,t)$ le champ électrique complexe de composantes $\mathcal{E}_x$ et $\mathcal{E}_y$ dans le plan $z = 0$:

$$\vec{\mathcal{E}}(0,t) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} e^{-i\omega t} \quad (\alpha \text{ et } \beta \text{ constantes complexes}) .$$

Le vecteur complexe de composantes α et β s'appelle **vecteur de Jones**. Par définition les composantes E_x et E_y du champ électrique réel $\vec{E}(0,t)$ sont obtenues en prenant les parties réelles de $\mathcal{E}_x$ et $\mathcal{E}_y$:

$$E_x = \Re e\big(\alpha e^{-i\omega t}\big) = \frac{\alpha e^{-i\omega t} + \overline{\alpha} e^{i\omega t}}{2} \quad ; \quad E_y = \Re e\big(\beta e^{-i\omega t}\big) = \frac{\beta e^{-i\omega t} + \overline{\beta} e^{i\omega t}}{2} .$$

L'extrémité de $\vec{E}(0,t)$ décrit une ellipse dans le plan (xOy) (**polarisation elliptique**). En effet sur l'expression

$$Z(t) = E_x + iE_y = \frac{\overline{\alpha} + i\overline{\beta}}{2} e^{i\omega t} + \frac{\alpha + i\beta}{2} e^{-i\omega t}$$

du nombre complexe associé au vecteur réel $\vec{E}$, on reconnait que $\vec{E}$ est la somme de deux vecteurs tournants. Les caractéristiques a, b et θ de l'ellipse ont été décrites à la section 2.3.4. En particulier la polarisation est circulaire gauche si $\alpha = -i\beta$, droite si $\alpha = i\beta$ et rectiligne si $|\alpha + i\beta| = |\overline{\alpha} + i\overline{\beta}|$. Inversement si on se donne des polarisations elliptiques gauche et droite (resp. une flèche et deux flèches sur la figure 33), de caractéristiques a, b et θ, les champs complexes $\vec{\mathcal{E}}_\pm(0,t)$ correspondant s'écrivent :

$$\vec{\mathcal{E}}_\pm(0,t) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ \pm ib \end{pmatrix} e^{-i(\omega t+\varphi)} .$$

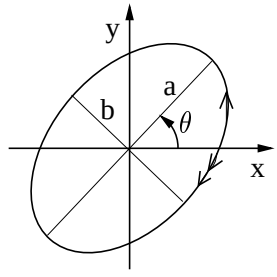


FIGURE 33

En effet $\Re e\, \vec{\mathcal{E}}_\pm(0,t)$ se déduit de la polarisation elliptique $a\,\cos(\omega t+\varphi)\,\hat{x} \pm b\,\sin(\omega t+\varphi)\,\hat{y}$ d'axes (x,y) par la rotation d'angle θ. La polarisation est circulaire si $a = b$ et rectiligne si $b = 0$.

▷ Effets de la propagation sur la polarisation

Ils s'obtiennent en observant que le champ en z n'est autre que le champ en $z = 0$ après propagation. Dans le cas d'un **milieu isotrope** on a

$$\vec{\mathcal{E}}(z,t) = \vec{\mathcal{E}}\big((0, t - \frac{z}{v}\big) = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} e^{-i\omega(t-\frac{z}{v})} \ ;$$

la trajectoire parcourue par l'extrémité du champ réel $\vec{E}(z,t)$ est donc identique à celle parcourue par l'extrémité de $\vec{E}(0,t)$, avec un retard de $\dfrac{z}{v}$, et la nature de la polarisation n'est donc pas modifiée.

Dans les **milieux anisotropes** on rencontre des situations où la vitesse de propagation n'est pas la même pour E_x et E_y :

$$\mathcal{E}_x(z,t) = \alpha\, e^{-i\omega(t-\frac{z}{v_x})} \quad ; \quad \mathcal{E}_y(z,t) = \beta\, e^{-i\omega(t-\frac{z}{v_y})} \ .$$

Dans ce cas la nature de la polarisation change au cours de la propagation. Pour une origine de temps bien choisie le champ réel $\vec{E}(z,t)$ a pour composantes :

$$E_x(t) = |\alpha|\, \cos\, \omega t \quad ; \quad E_y(t) = |\beta|\, \cos\,(\omega t + \varphi) \quad ; \quad \varphi = \arg\, \alpha - \arg\, \beta + \omega z\big(\frac{1}{v_x} - \frac{1}{v_y}\big) \ .$$

Les trajectoires possibles pour son extrémité sont des ellipses dépendant de φ, inscrites dans un rectangle de côtés $2|\alpha|$, $2|\beta|$ (figure 34). Leurs "inclinaison" et sens de parcours sont déterminés respectivement par la valeur de $E_y(0)$ et le signe de $\dfrac{\mathrm{d}E_y}{\mathrm{d}t}(0)$.

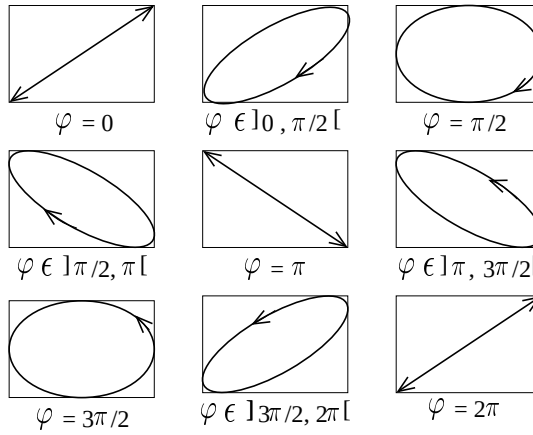


FIGURE 34

Pour les **milieux gyrotropes** ce sont les polarisations circulaires gauche et droite qui ont des vitesses de propagation bien définies et différentes :

$$\vec{\mathcal{E}}_+(z,t) \propto \begin{pmatrix}1\\ i\end{pmatrix} e^{-i\omega(t-\frac{z}{v_+})} \quad ; \quad \vec{\mathcal{E}}_-(z,t) \propto \begin{pmatrix}1\\ -i\end{pmatrix} e^{-i\omega(t-\frac{z}{v_-})} \ .$$

Pour voir comment évolue une polarisation quelconque en fonction de z, le plus simple est de dessiner, au même instant, les vecteurs tournants qui la composent dans le plan $z = 0$ (lorsque ces vecteurs sont parallèles) et dans un plan $z > 0$ (figure 35).

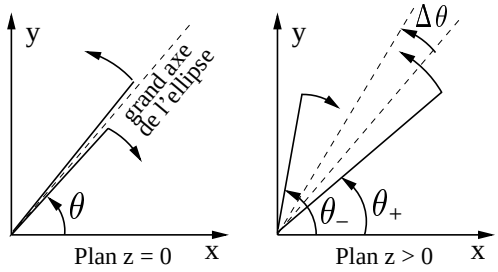


FIGURE 35

Leurs modules étant inchangés, la polarisation a les mêmes paramètres a et b et le même sens de rotation. Cependant, comme dans le plan $z > 0$ ils sont en retard respectivement de $\dfrac{z}{v_+}$ et de $\dfrac{z}{v_-}$, leurs positions sont $\theta_\pm = \theta \mp \dfrac{\omega z}{v_\pm}$; les axes de l'ellipse ont donc tourné de $\Delta\theta = \dfrac{\theta_+ + \theta_-}{2} - \theta = \dfrac{1}{2}\,\omega z \left(\dfrac{1}{v_-} - \dfrac{1}{v_+}\right)$.

▷ Interférences de deux ondes polarisées monochromatiques

L'interférence de deux ondes cohérentes 1 et 2, de même pulsation mais de polarisations différentes, entre lesquelles on introduit un déphasage φ, conduit à une intensité (somme des contributions des deux composantes du champ électrique) :

$$I = \frac{1}{2}\,|\alpha_1 + \alpha_2\, e^{-i\varphi}|^2 + \frac{1}{2}\,|\beta_1 + \beta_2\, e^{-i\varphi}|^2 \quad = I_0\big(1 + C\,\cos(\varphi - \varphi_0)\big).$$

Dans cette expression $\varphi_0 = \arg(\overline{\alpha_1}\alpha_2 + \overline{\beta_1}\beta_2)$, $I_0 = \dfrac{1}{2}\,(|\alpha_1|^2 + |\beta_1|^2 + |\alpha_2|^2 + |\beta_2|^2) = I_1 + I_2$ et $C = \dfrac{|\overline{\alpha_1}\alpha_2 + \overline{\beta_1}\beta_2|}{I_1 + I_2}$ est le contraste. On voit que le terme d'interférence disparaît si $\overline{\alpha_1}\alpha_2 + \overline{\beta_1}\beta_2 = 0$; cela correspond à deux polarisations elliptiques gauche et droite de grands axes perpendiculaires, par exemple : $(\alpha_1, \beta_1) = (a, ib)$ et $(\alpha_2, \beta_2) = (b, -ia)$.

▷ Ondes polarisées non monochromatiques

Dans ce cas $\alpha(t)$ et $\beta(t)$ sont des fonctions du temps et la polarisation elliptique fluctue. Les intensités lumineuses pouvant être mesurées expérimentalement font intervenir la **matrice de Jones** :

$$\mathbf{J} = \begin{pmatrix} <|\alpha|^2> & <\alpha\overline{\beta}> \\ <\overline{\alpha}\beta> & <|\beta|^2> \end{pmatrix}.$$

EXEMPLES. Un polariseur ne laissant passer que la composante du champ parallèle à une direction fixe θ_0 mesure une intensité proportionnelle à $<(E_x\,\cos\theta_0+E_y\,\sin\theta_0)^2>$ $=\dfrac{1}{2}\,<|\alpha\,\cos\,\theta_0+\beta\,\sin\,\theta_0|^2>$, soit :

$$I=\frac{1}{2}\,(<|\alpha|^2>\cos^2\theta_0+2\Re e<\overline{\alpha}\beta>\cos\theta_0\,\sin\,\theta_0+<|\beta|^2>\sin^2\theta_0)\ .$$

Un système ne laissant passer que les polarisations circulaires directes mesure (*cf.* expression de $Z(t)$ ci-dessus) :

$$I=<|\frac{\overline{\alpha}+i\overline{\beta}}{2}|^2>=\frac{1}{4}\,\left(<|\alpha|^2>+<|\beta|^2>+2\Im m\,<\overline{\alpha}\beta>\right)\ .$$

Propriétés de J. **J** est une matrice hermitienne positive :

$$\mathrm{tr}\,\mathbf{J}=<|\alpha|^2>+<|\beta|^2>>0\quad ;\quad \det\mathbf{J}=<|\alpha|^2><|\beta|^2>-|<\overline{\alpha}\beta>|^2\ \geq 0\ .$$

$\det\mathbf{J}\geq 0$ est une conséquence de l'inégalité de Schwarz $<|\lambda\alpha+\beta|^2>\geq 0$ quel que soit $\lambda\in\mathbb{C}$. En effet $<|\lambda\alpha+\beta|^2>=<|\alpha|^2>|\lambda|^2+2\Re e(\lambda<\alpha\overline{\beta}>)+<|\beta|^2>\geq 0$ et en particulier : $<|\alpha|^2>|\lambda|^2+2|\lambda||<\alpha\overline{\beta}>|+<|\beta|^2>\geq 0$.
Le cas où **J** est multiple de l'identité correspond à une **lumière naturelle** c'est-à-dire **non polarisée** ; celui où $\det\mathbf{J}=0$ correspond à une **lumière complètement polarisée** (comme pour α et β constants).

Décompositions de J

1) On vérifie aisément par identification qu'il est toujours possible de décomposer de manière unique **J** sous la forme

$$\mathbf{J}=\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}+\begin{pmatrix}<|\alpha|^2>-\lambda & <\overline{\alpha}\beta>\\ <\alpha\overline{\beta}> & <|\beta|^2>-\lambda\end{pmatrix}=\mathbf{J}_1+\mathbf{J}_2\ ,$$

avec $0<\lambda\leq\ (<|\alpha|^2>\ \text{et}\ <|\beta|^2>)$ et $\det\mathbf{J}_2=0$; on interprète alors le champ électrique comme la somme des champs (incohérents entre eux) des lumières ci-dessus.

2) Une autre décomposition possible est $\mathbf{J}=\lambda_1\,\mathbf{P}_1+\lambda_2\,\mathbf{P}_2$ où les $\mathbf{P}_i$ sont les projecteurs sur les vecteurs propres orthogonaux de **J** (*cf.* section 4.2.3) ; le champ est alors la somme de deux polarisations orthogonales incohérentes entre elles. $\rho=(\mathrm{tr}\,\mathbf{J})^{-1}\,\mathbf{J}$ est la **matrice densité** des photons associés (*cf.* section 4.4.1).

Chapitre 3

Espace ; symétries ; calcul vectoriel

La géométrie, autrefois présentée comme "art de raisonner juste sur des figures fausses", s'est très souvent développée au cours de l'histoire en relation directe avec l'étude et la représentation de l'"espace physique" : **géométrie euclidienne** (Euclide 400 A.C.) vérifiée "expérimentalement" avec une très grande précision ; géométrie projective introduite à la Renaissance en relation avec les problèmes de représentation de la perspective (Desargues, Pascal $\simeq$ 1640) ; géométries non euclidiennes ($\simeq$ 1830) dont les auteurs, notamment Gauss et Lobatchevski, ont cherché à tester les conséquences observationnelles ; géométrie des espaces métriques dont Riemann ($\simeq$ 1880) a pensé pouvoir associer la courbure à la gravitation *etc.* Le concept important à la base de toutes ces géométries, et d'autres telles que la géométrie symplectique (pour l'espace de phase en mécanique) ou la géométrie des espaces fibrés (pour décrire les interactions fondamentales), est celui de **groupe de symétrie**. Le mathématicien Klein a été le premier dans son programme d'Erlangen (1872) à insister sur l'idée qu'un espace géométrique est défini par une dimension et l'action d'un groupe de symétrie, et à avoir défini les **propriétés géométriques** par leur invariance vis-à-vis de ce groupe. Par exemple le plan "plat" et la sphère "courbe" sont deux espaces de dimension 2 **homogènes** (pas de point privilégié) et **isotropes** (pas de direction privilégiée en chaque point), mais dont les groupes de symétrie, respectivement le groupe euclidien et le groupe des rotations, diffèrent.

En physique l'espace n'est défini que par rapport à un référentiel. Depuis Galilée (1632) on sait que les référentiels d'inertie se caractérisent par l'**invariance des lois physiques** vis-à-vis des translations et rotations des coordonnées spatiales dans le référentiel (homogénéité et isotropie de l'espace), des translations de temps (oubliées par Galilée), et des mises en mouvement à vitesse uniforme de ces référentiels. Einstein (1905) a simplement étendu ce "principe de relativité" aux lois de l'électromagnétisme. La comparaison avec les géométries montre qu'il existe donc une correspondance entre groupe de symétrie d'un espace et **groupe de relativité** de l'espace temps d'une part, et entre propriétés géométriques et lois physiques d'autre part. Physique galiléenne et physique einsteinienne diffèrent par leur groupe de symétrie.

Après une présentation générale des notions de symétrie et d'invariance en physique, nous nous intéressons principalement dans ce chapitre aux symétries spatiales (translations, rotations et symétries ponctuelles). Il est en effet indispensable de connaître leur action non seulement sur l'espace (objet de la géométrie euclidienne), mais aussi sur les grandeurs physiques qu'elles permettent de classer en scalaires, vecteurs... La description mathématique de cette action repose pour une part importante sur les outils du **calcul vectoriel** (produits scalaire, vectoriel, mixte) mis au point vers 1880 par les mathématiciens et les physiciens (Grassman, Hamilton, Gibbs, Maxwell...). Elle permet, lorsqu'un système physique possède des symétries spatiales particulières, de prédire certaines de ses propriétés grâce à des "**arguments de symétrie**". Des compléments sur les symétries spatiales (quadrupoles, spineurs...) ou autres (Lorentz en particulier) sont donnés au chapitre 4.

3.1 SYMÉTRIE, INVARIANCE ET RELATIVITÉ

3.1.1 Groupes de symétrie et invariance

▷ Groupes et représentations

Un groupe G (Galois 1832) est un ensemble d'éléments g muni d'un produit associatif $(g_3(g_2g_1) = (g_3g_2)g_1)$ avec un élément neutre e $(eg = ge = g)$, chaque élément g ayant un inverse g^{-1} $(gg^{-1} = g^{-1}g = e)$. Il est commutatif si $g_1g_2 = g_2g_1$ quels que soient $g_1, g_2 \in G$; c'est en particulier le cas de tous les groupes continus à un seul paramètre, par exemple les rotations autour d'un axe fixe.

Un groupe est donc un ensemble abstrait défini par sa loi de multiplication ; mais en pratique il "agit" sur des objets géométriques (points, vecteurs...) ou sur des systèmes et des grandeurs physiques (dipole, champ électrique...) caractérisés par des paramètres $\underline{x}$ (coordonnées de points, composantes de vecteurs...). Cette action notée un peu abusivement $\underline{x} \to g\underline{x}$, qui vérifie $e\underline{x} = \underline{x}$ et $g_2(g_1\underline{x}) = (g_2g_1)\underline{x}$, décrit une représentation du groupe G dans l'espace X des paramètres. En général les espaces X correspondant aux différentes grandeurs physiques ont une structure d'espace vectoriel ; les **représentations linéaires**, qui vérifient par définition $g(\lambda_1\underline{x}_1 + \lambda_2\underline{x}_2) = \lambda_1\,(g\underline{x}_1) + \lambda_2\,(g\underline{x}_2)$, jouent alors un rôle important pour la classification de ces grandeurs.

▷ Symétries et invariance

Un objet (ou système physique) $\underline{x}$ est invariant par rapport à une transformation g si $g\underline{x} = \underline{x}$ (paramètres inchangés). L'ensemble des transformations qui laissent $\underline{x}$ invariant constitue son **groupe de symétrie**. Mais en géométrie comme en physique le concept de symétrie est plus profond, car il ne concerne pas simplement les objets pris individuellement mais plus généralement leurs relations. Une relation est dite invariante si, étant vraie pour des objets, elle l'est aussi pour leurs transformés ; symboliquement :

$$f(\underline{x}_1, \underline{x}_2 \cdots) = 0 \Longleftrightarrow f(g\underline{x}_1, g\underline{x}_2 \cdots) = 0$$

pour tout $g \in G$. G est alors un groupe de symétrie pour cette relation. Le groupe de symétrie des propriétés géométriques ou des lois physiques est toujours plus grand que celui de tel ou tel objet particulier. Par exemple la propriété pour l'ensemble de deux droites du plan d'être orthogonales est laissée invariante par toutes les translations et

toutes les rotations, mais seules les rotations de $\frac{n\pi}{2}$ autour de leur point d'intersection laissent cet ensemble invariant. De même la loi selon laquelle un point M massif sur une planète sphérique est soumis à une force dirigée vers le centre O est laissée invariante par toutes les translations et rotations de l'ensemble "masse-planète", alors que ce système n'est laissé invariant que par les rotations autour de l'axe OM.

▷ Points de vue actif et passif

Les transformations $\underline{x} \to \mathrm{g}\underline{x}$ qui font passer d'un objet paramétré par $\underline{x}$ dans un système de coordonnées $\mathcal{R}$ à l'objet paramétré par $\mathrm{g}\underline{x}$ (toujours dans $\mathcal{R}$) sont dites actives. On peut inversement, point de vue passif, effectuer l'opération g sur le système de coordonnées ; $\mathcal{R}$ devient alors $\mathcal{R}$' et le même objet paramétré par $\underline{x}$ dans $\mathcal{R}$ l'est par $\underline{x}' = \mathrm{g}^{-1}\underline{x}$ dans $\mathcal{R}$'.
DÉMONSTRATION : si on applique g à la fois au système de coordonnées $\mathcal{R}$ et aux objets les coordonnées ne changent pas (figure 1 : $\mathcal{R}$ est arbitraire et g est une rotation). Donc $(\mathrm{g}\underline{x})' = \underline{x}$ pour tout $\underline{x}$ et $\underline{x}' = (\mathrm{g}\mathrm{g}^{-1}\underline{x})' = \mathrm{g}^{-1}\underline{x}$. Les deux points de vue sont équivalents.

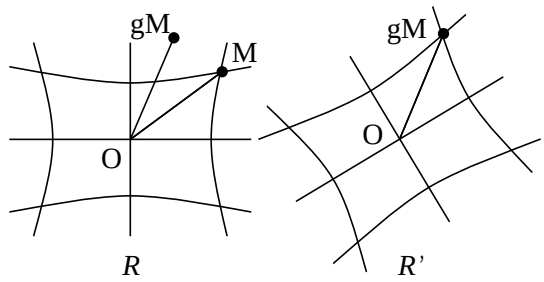


FIGURE 1

Une propriété (ou loi) qui s'exprime par la relation $f(\underline{x}_1, \underline{x}_2 \cdots) = 0$ dans $\mathcal{R}$ sera dite invariante si elle s'exprime sous forme identique à la précécente $f(\underline{x}'_1, \underline{x}'_2 \cdots) = 0$ dans tous les systèmes $\mathcal{R}$' déduits de $\mathcal{R}$ par les transformations $\mathrm{g} \in G$. (Remarquons l'analogie avec les changements d'unités étudiés à la section 1.3.1.)

3.1.2 Le groupe de symétrie de la physique

Dans ce paragraphe nous adoptons exceptionnellement le point de vue passif et nous désignons par $\mathcal{R}$ un système de coordonnées (ou référentiel) muni d'une origine spatio-temporelle, ce qui permet d'affecter à tout évènement une coordonnée temporelle t et trois coordonnées spatiales x, y, z (relatives à un trièdre orthonormé). Considérons deux évènements repérés dans $\mathcal{R}$ par (t_i, x_i, y_i, z_i) $(i = 1, 2)$, et soient $T = t_2 - t_1$, $X = x_2 - x_1$, $Y = y_2 - y_1$ et $Z = z_2 - z_1$ ($\overrightarrow{R} = \overrightarrow{r}_2 - \overrightarrow{r}_1$) les composantes de l'**intervalle d'espace-temps** entre ces évènements. Les changements de référentiels inertiels sont caractérisés par les deux propriétés :

(1) préservation de l'addition des intervalles (2) invariance de $c^2T^2 - \overrightarrow{R}^2$.

(1) traduit l'homogénéité d'un espace temps "plat" et implique que les transformations sur T, X, Y et Z sont linéaires. (2) qui s'écrit $c^2T^2 - \overrightarrow{R}^2 = c^2T'^2 - \overrightarrow{R}'^2$ a été intro-

duite historiquement pour rendre compte de l'invariance de la vitesse c de la lumière : $\dfrac{|\overrightarrow{R}|}{T} = c \Longleftrightarrow \dfrac{|\overrightarrow{R}'|}{T'} = c$. On peut montrer que le groupe des transformations continues compatibles avec ces deux propriétés s'obtient par composition des trois types suivants de transformations :

- **Les translations d'espace-temps** qui ne changent pas les composantes d'un intervalle ($T' = T$, $\overrightarrow{R}' = \overrightarrow{R}$) mais translatent les coordonnées d'un évènement :

$$t' = t - t_0 \quad , \quad x' = x - x_0 \quad , \quad y' = y - y_0 \quad , \quad z' = z - z_0 \ .$$

Elles dépendent de quatre paramètres, les coordonnées t_0, x_0, y_0 et z_0 dans $\mathcal{R}$ de l'évènement choisi comme origine dans $\mathcal{R}'$.

- **Les rotations** qui vérifient $T' = T$ et $|\overrightarrow{R}'| = |\overrightarrow{R}|$; une rotation dépend de trois paramètres (l'angle et l'axe de rotation, ce dernier étant spécifié par deux angles). Par exemple pour une rotation d'axe Oz (figure 2) :

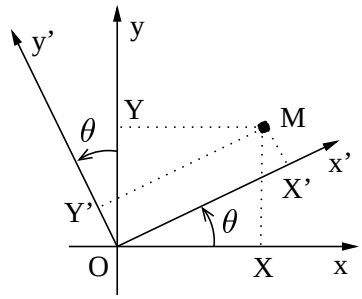


FIGURE 2

$$T' = T \quad ; \quad X' = X\cos\theta + Y\sin\theta \quad ; \quad Y' = -X\sin\theta + Y\cos\theta \quad ; \quad Z' = Z \ .$$

- **Les transformations de Lorentz pures** (sans rotation) caractérisées par la vitesse de translation uniforme $\overrightarrow{V}$ du référentiel $\mathcal{R}'$ par rapport au référentiel $\mathcal{R}$; elles dépendent donc de trois paramètres. Elles concernent à la fois T et la composante de $\overrightarrow{R}$ parallèle à la vitesse de translation, par exemple :

$$cT' = cT\cosh\varphi - X\sinh\varphi \quad , \quad X' = X\cosh\varphi - cT\sinh\varphi \quad ; \quad Y' = Y \quad , \quad Z' = Z;$$

l'invariance de $c^2T^2 - X^2$ résulte de $\cosh^2\varphi - \sinh^2\varphi = 1$. Interprétation physique : le référentiel $\mathcal{R}'$ se déplace à la vitesse $V = c\tanh\varphi$ (φ **rapidité**) dans la direction x par rapport à $\mathcal{R}$; en effet en posant $X' = Y' = Z' = 0$, c'est-à-dire en considérant tous les évènements attachés à un point fixe de $\mathcal{R}'$, les relations $X = c\tanh\varphi\, T$ et $Y = Z = 0$ montrent que leurs positions dans $\mathcal{R}$ se déplacent dans la direction x à la vitesse V. Comme $1 - \tanh^2\varphi = \cosh^{-2}\varphi$ la transformation s'écrit aussi :

$$T' = \gamma\left(T - \frac{VX}{c^2}\right) \quad ; \quad X' = \gamma\,(X - VT) \qquad \text{avec } \gamma = \cosh\varphi = \left(1 - \frac{V^2}{c^2}\right)^{-\frac{1}{2}}.$$

REMARQUE (figure 3). A un instant donné dans $\mathcal{R}$ ($T = 0$), des points fixes dans $\mathcal{R}'$ espacés dans $\mathcal{R}'$ de X', Y', Z' sont représentés dans $\mathcal{R}$ avec un espacement $X = \gamma^{-1}X'$, $Y = Y'$, $Z = Z'$: **contraction "apparente"** dans la direction x, négligeable si $\dfrac{V^2}{c^2} \ll 1$.

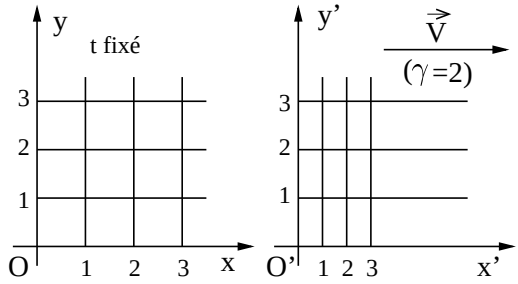


FIGURE 3

▷ Approximation galiléenne

Dans la double limite $\dfrac{X}{cT}$ et $\dfrac{V}{c} \ll 1$, la transformation devient $T' = T$ et $X' = X - VT$. Plus généralement on appelle physique galiléenne l'ensemble des lois laissées invariantes par les translations, les rotations, et les **transformations de Galilée** pures

$$T' = T \quad , \quad \overrightarrow{R}' = \overrightarrow{R} - \overrightarrow{V}T \ .$$

REMARQUE 1. Aux symétries "continues" ci-dessus peuvent s'ajouter des **symétries discrètes** qui correspondent à des transformations qui ne peuvent pas être atteintes continûment à partir de l'identité. Il en est ainsi de l'opération de **renversement du temps** ($T' = -T$, $\overrightarrow{R}' = \overrightarrow{R}$) qui est une symétrie pour les phénomènes réversibles, ou de la **parité** ($T' = T$, $\overrightarrow{R}' = -\overrightarrow{R}$) qui est une symétrie pour toutes les interactions sauf l'interaction faible (*cf.* section 3.2.1).

REMARQUE 2. Vocabulaire : on dit de façon incorrecte que les lois physiques sont "**relativistes**". Elles sont en fait "**invariantes**" (même formulation dans tous les référentiels inertiels). Ce qui est relatif, c'est la valeur des grandeurs, à commencer par la grandeur "espace" $\overrightarrow{R}$ qui, même en physique galiléenne (si $T \neq 0$), dépend du référentiel. De ce point de vue il n'y a pas de différence avec la géométrie : une égalité entre vecteurs $\overrightarrow{V} = \overrightarrow{W}$ se formule de la même manière dans deux bases différentes : $V_i = W_i \Longleftrightarrow V_i' = W_i'$, mais $V_i' \neq V_i$ et $W_i' \neq W_i$.

3.1.3 Symétries spatiales (présentation « expérimentale »)

Les symétries spatiales continues correspondent aux déplacements des objets dans l'espace. Comme on le sait, un déplacement conserve les distances entre points et, en général, "tourne" les directions (en conservant les angles). Un **déplacement** général s'obtient en composant des translations et des rotations.

▷ Translations

Les translations conservent les directions ; elles forment un groupe commutatif dont la loi de composition correspond à l'addition vectorielle (figure 4a). La translation $\mathrm{T}_{\overrightarrow{a}}$ de vecteur $\overrightarrow{a}$ amène tout point M en M' tel que $\overrightarrow{MM'} = \overrightarrow{a}$.

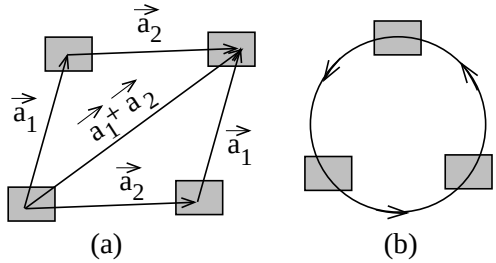


FIGURE 4

REMARQUE. Lorsque le vecteur $\overrightarrow{a}(t)$ dépend du temps, on parle de **mouvement de translation** (figure 4b). Par exemple le référentiel géocentrique ayant pour origine le centre T de la Terre et ses axes parallèles à ceux du référentiel de Copernic centré sur le Soleil S effectue un mouvement de translation autour du Soleil, et non de rotation bien que le vecteur $\overrightarrow{ST}$ tourne.

▷ Rotations de centre O

Ces rotations, définies par un axe passant par O et un angle, "tournent" les directions. Elles forment un groupe mais ne commutent pas à l'exception du cas de rotations de même axe ou du cas de rotations d'angle π et d'axes orthogonaux. La loi de composition générale est donnée à la section 4.2.3.

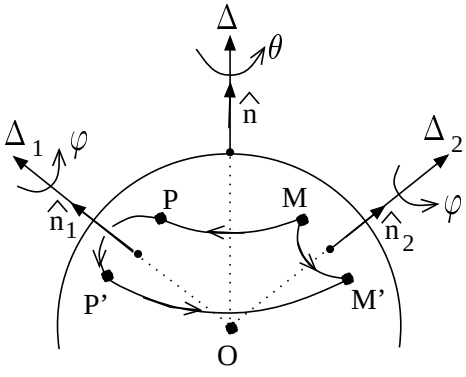


FIGURE 5

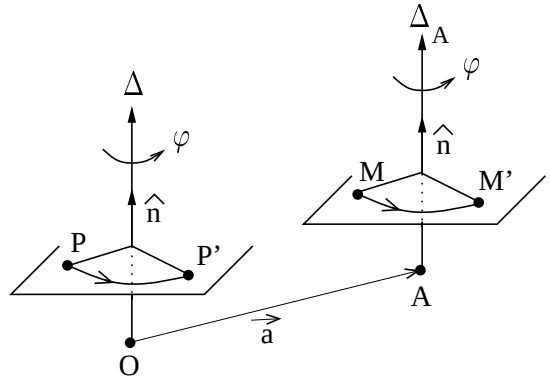


FIGURE 6

JUSTIFICATION (figure 5). Considérons l'expression $\mathrm{R}_O(\hat{n}_2,\varphi) = \mathrm{R}_O(\hat{n},\theta)\,\mathrm{R}_O(\hat{n}_1,\varphi)\,\mathrm{R}_O^{-1}(\hat{n},\theta)$ qui relie deux rotations de même angle φ dont les axes, de vecteurs directeurs $\hat{n}_1$ et $\hat{n}_2$, se déduisent l'un de l'autre par la rotation $\mathrm{R}_O(\hat{n},\theta)$ amenant $\hat{n}_1$ sur $\hat{n}_2$. Les rotations a priori quelconques $\mathrm{R}_O(\hat{n},\theta)$ et $\mathrm{R}_O(\hat{n}_1,\varphi)$ commutent si $\mathrm{R}_O(\hat{n},\theta)\,\mathrm{R}_O(\hat{n}_1,\varphi)\,(= \mathrm{R}_O(\hat{n}_2,\varphi)\,\mathrm{R}_O(\hat{n},\theta)) = \mathrm{R}_O(\hat{n}_1,\varphi)\,\mathrm{R}_O(\hat{n},\theta)$, donc si $\mathrm{R}_O(\hat{n}_1,\varphi) = \mathrm{R}_O(\hat{n}_2,\varphi)$). Cette relation entraîne soit $\hat{n}_1 = \hat{n}_2$ si φ est quelconque, soit $\hat{n}_2 = -\hat{n}_1$ et $\varphi = \pi$, ce qui conduit aux deux cas cités (soit $\hat{n} = \hat{n}_1$, soit $\hat{n} \perp \hat{n}_1$ et $\theta = \varphi = \pi$).

REMARQUE (figure 6). Des rotations d'axes parallèles et de même angle, mais de centres A et O différents, se déduisent par la relation :

$$\mathrm{R}_A(\hat{n},\varphi) = \mathrm{T}_{\overrightarrow{a}}\mathrm{R}_O(\hat{n},\varphi)\mathrm{T}_{-\overrightarrow{a}} \qquad (\overrightarrow{a} = \overrightarrow{OA}) \ .$$

▷ Vissages

La relation ci-dessus montre qu'une translation $\mathrm{T}_{\overrightarrow{a}}$ ne commute avec une rotation $\mathrm{R}_O(\hat{n},\varphi)$ que si $\overrightarrow{a}$ est parallèle à l'axe de rotation $\hat{n}$ (ce qui assure $\mathrm{R}_O(\hat{n},\varphi) = \mathrm{R}_A(\hat{n},\varphi)$ puisque O et A sont alors sur le même axe). La composition des deux transformations, lorsqu'elles commutent, est un vissage ; sur la figure 7, l'**hélice** γ est le lieu des points qui se déduisent de M par les vissages $\lambda\overrightarrow{a}$, $\lambda\varphi$ $(0 < \lambda < 1)$.

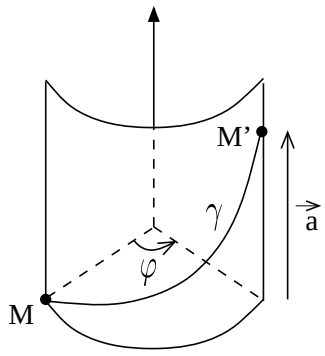


FIGURE 7

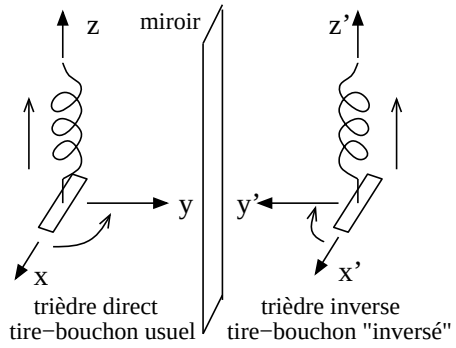


FIGURE 8

Un exemple est donné par le tire-bouchon (usuel) dont le sens de déplacement, lié au sens de rotation, sert à définir les trièdes directs par la fameuse "**règle du tire-bouchon**" (figure 8). Tout **déplacement** se ramène à un vissage.

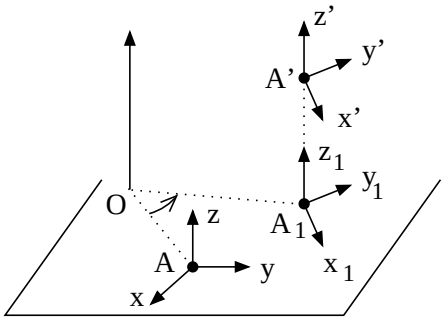


FIGURE 9

JUSTIFICATION (figure 9) : restreinte aux vecteurs unitaires $\hat{n}$ qui caractérisent les directions (vecteurs "libres"), l'effet d'un déplacement se résume à une rotation. Soit z la direction de son axe. Un trièdre $Axyz$ est alors déplacé en $A'x'y'z'$ tel que les axes Az et $A'z'$ sont parallèles. La figure montre que ce déplacement s'obtient en composant un **déplacement plan** amenant $Axyz$ en $A_1x_1y_1z_1$ (A_1, x_1, y_1 projection de A', x', y' sur le plan Axy), et une translation $\overrightarrow{A_1A'}$ parallèle à z. Le déplacement plan est

une rotation dont l'axe parallèle à z (axe du vissage) coupe le plan en O, **centre de rotation** situé sur la médiatrice de AA_1 et tel que l'angle $(\overrightarrow{OA}, \overrightarrow{OA_1})$ est celui dont ont tourné les axes Ax et Ay. Si ces axes n'ont pas tourné le déplacement est simplement une translation.

Mouvement d'un solide. C'est une succession de vissages élémentaires reliant ses positions aux instants t et $t + \mathrm{d}t$. Chaque **vissage élémentaire** est caractérisé par un axe (dont la position et la direction $\hat{n}(t)$ dépendent du temps), une rotation infinitésimale d'angle $\omega(t)\,\mathrm{d}t$ autour de cet axe et une translation infinitésimale de vecteur $\overrightarrow{v}(t)\,\mathrm{d}t$ parallèle à $\hat{n}(t)$ (*cf.* section 3.3.3).

▷ Symétries discrètes (figure 10)

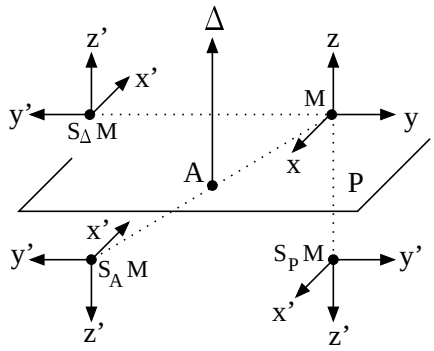


FIGURE 10

Elles conservent aussi les distances et les angles (en valeurs absolues). On distingue les symétries S_A par rapport à un centre A ($\overrightarrow{AM} \to \overrightarrow{AM'} = -\overrightarrow{AM}$) , S_Δ par rapport à une droite Δ (équivalente à une rotation de π autour de Δ), et S_P par rapport à un plan P. Dans tous les cas S^2 est la transformation identité. Si $\Delta \perp P$ coupe P en A, S_P diffère de S_A par S_Δ. Noter que S_A et S_P, mais pas S_Δ, changent le caractère (direct ou indirect) d'un trièdre.

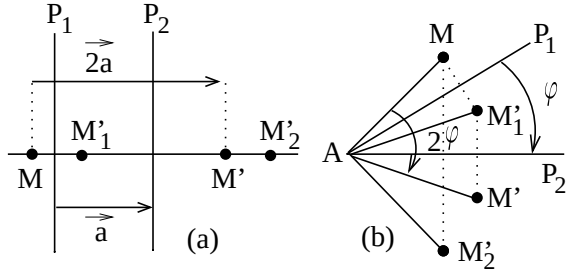


FIGURE 11

Les **lois de composition** de deux symétries par rapport à des plans P_1 et P_2 qui, soit sont parallèles (ils se déduisent alors l'un de l'autre par la translation de vecteur $\overrightarrow{a} \perp P_1$ et P_2), soit se coupent selon l'axe $(A, \hat{n})$ (la rotation $\mathrm{R}_A(\hat{n}, \varphi)$ transformant P_1 en P_2),

vérifient respectivement (figures 11a,b, où $M'_1 = \mathrm{S}_{P_1} M$ et $M' = \mathrm{S}_{P_2}\mathrm{S}_{P_1} M$) :

$$\mathrm{S}_{P_2}\mathrm{S}_{P_1} = \mathrm{T}_{2\overrightarrow{a}} \quad \text{et} \quad \mathrm{S}_{P_2}\mathrm{S}_{P_1} = \mathrm{R}_A(\hat{n}, 2\varphi) \ .$$

Ecrites $\mathrm{S}_{P_2} = \mathrm{T}_{2\overrightarrow{a}}\mathrm{S}_{P_1}$ et $\mathrm{S}_{P_2} = \mathrm{R}_A(\hat{n}, 2\varphi)\,\mathrm{S}_{P_1}$, ces relations montrent que les symétriques M'_1 et M'_2 d'un point M par rapport aux plans P_1 et P_2 se déduisent l'un de l'autre, soit par la translation $\mathrm{T}_{2\overrightarrow{a}}$, soit par la rotation $\mathrm{R}_A(\hat{n}, 2\varphi)$.

EXEMPLES. Soient S_1 et S_2 les images d'une source S données par des **miroirs parallèles ou "en coin"**, et γ_1 et γ_2 les deux rayons issus d'un même rayon incident γ (figures 12a,b). Les intersections L de γ_1 et γ_2 donnent le lieu de localisation des franges d'interférences (*cf.* section 3.2.3). Pour la figure 12a L est à l'infini. Pour la figure 12b (dans le plan de figure), L est sur le cercle circonscrit au triangle S_1OS_2 car γ_1 et γ_2 font entre eux un angle 2φ comme OS_1 et OS_2 ; ce cercle devient la droite OP lorsque S est à l'infini dans la direction i. La figure 12c illustre le principe du **cataphote** : $\mathrm{S}_{P_1}\mathrm{S}_{P_2}\mathrm{S}_{P_3} = \mathrm{S}_O$ quel que soit l'ordre des réflexions sur les trois plans perpendiculaires deux à deux.

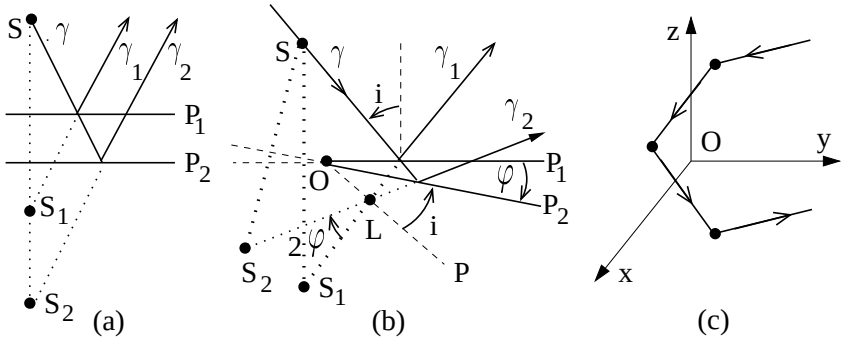


FIGURE 12

3.1.4 Transformation des grandeurs et des champs physiques

Les opérations de symétrie ne portent pas uniquement sur des objets géométriques "ensembles de points", mais aussi sur des objets physiques auxquels sont associées des grandeurs. Lorsque l'objet est déplacé (ou subit par la pensée une symétrie ponctuelle), ces grandeurs peuvent changer. Toutes les grandeurs classiques sont laissées inchangées par les translations. (Voir remarque section 3.4.2. pour un état quantique.) Par contre vis-à-vis des rotations et des symétries ponctuelles, on distingue les **grandeurs scalaires** laissées inchangées (la masse, la charge, la température, *etc.*), les **grandeurs vectorielles** qui se transforment comme les vecteurs de la géométrie (vitesse, moment dipolaire, *etc.*), et d'autres (spineurs, quadrupoles...) étudiées au chapitre 4. On verra aussi à la section 3.2.1 que, vis-à-vis des symétries discrètes, il faut distinguer les grandeurs scalaires et **pseudoscalaires** de même que les grandeurs vectorielles et **pseudovectorielles**. Dans tous les cas, l'invariance des lois physiques entraîne que les égalités ne peuvent avoir lieu qu'entre grandeurs ayant les mêmes lois de transformation.

Ces opérations de symétrie portent aussi sur les champs relatifs aux objets physiques (densité volumique de charges par exemple), ou créés par eux (champ électrique), qui sont des grandeurs dont la valeur dépend du point d'espace où elle est mesurée.

La figure 13 illustre, sur l'exemple du potentiel $V(\overrightarrow{r})$ (**champ scalaire**) et du champ électrique $\overrightarrow{E}(\overrightarrow{r})$ (**champ vectoriel**) d'un dipole, le fait très important que dans une symétrie la valeur du champ transformé au point transformé s'obtient en appliquant la transformation à la valeur du champ initial au point initial, ce qu'on peut écrire symboliquement :

$$f'(\mathrm{g}\overrightarrow{r}) = \mathrm{g}\, f(\overrightarrow{r}) \ .$$

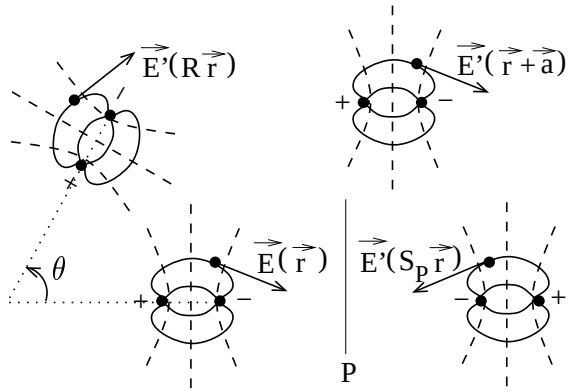


FIGURE 13

ATTENTION : g dans le membre de gauche désigne l'action de la symétrie sur le vecteur position, tandis que dans le membre de droite il désigne son action sur la valeur $f(\overrightarrow{r})$ de la grandeur f au point $\overrightarrow{r}$, action qui dépend de la nature (scalaire, vectorielle ou autre) de f. Ainsi pour une **translation**, $f'(\overrightarrow{r}+\overrightarrow{a}) = f(\overrightarrow{r})$ quelle que soit la nature de la grandeur f ; pour une **rotation**, $f'(\mathrm{R}\overrightarrow{r}) = f(\overrightarrow{r})$ pour un champ scalaire et $\overrightarrow{f}'(\mathrm{R}\overrightarrow{r}) = \mathrm{R}\,\overrightarrow{f}(\overrightarrow{r})$ pour un champ vectoriel. La connaissance de ces lois de transformation est essentielle pour comprendre l'application des arguments de symétrie aux champs. La poursuite de cette discussion sur les symétries est faite à la section 3.4.1.

3.2 CALCUL VECTORIEL ; APPLICATIONS

3.2.1 Produit scalaire, produit vectoriel et produit mixte ; pseudovecteurs

▷ Définitions et propriétés

Ces produits, qui généralisent $\overrightarrow{V}_1 \cdot \overrightarrow{V}_2$ et $\sigma(\overrightarrow{V}_1, \overrightarrow{V}_2)$ vus à la section 2.2.1, sont les quantités multilinéaires (*i.e.* linéaires par rapport à chacun des facteurs) définies dans une base orthonormée directe de vecteurs unitaires $\hat{x}$, $\hat{y}$ et $\hat{z}$ par :

$$\overrightarrow{V}_1 \cdot \overrightarrow{V}_2 = X_1X_2 + Y_1Y_2 + Z_1Z_2 = \overrightarrow{V}_2 \cdot \overrightarrow{V}_1 \ ,$$

$$\overrightarrow{V}_1 \wedge \overrightarrow{V}_2 = (Y_1Z_2 - Y_2Z_1)\hat{x} + (Z_1X_2 - Z_2X_1)\hat{y} + (X_1Y_2 - X_2Y_1)\hat{z} = -\overrightarrow{V}_2 \wedge \overrightarrow{V}_1 \ ,$$

$$\begin{aligned}(\overrightarrow{V}_1, \overrightarrow{V}_2, \overrightarrow{V}_3) &= (\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \cdot \overrightarrow{V}_3 \\ &= (Y_1Z_2 - Y_2Z_1)X_3 + (Z_1X_2 - Z_2X_1)Y_3 + (X_1Y_2 - X_2Y_1)Z_3 \\ &= (\overrightarrow{V}_2, \overrightarrow{V}_3, \overrightarrow{V}_1) = (\overrightarrow{V}_3, \overrightarrow{V}_1, \overrightarrow{V}_2) \quad \text{(permutations circulaires)} ;\end{aligned}$$

X_i, Y_i et Z_i sont les composantes des vecteurs $\overrightarrow{V}_i = X_i\hat{x} + Y_i\hat{y} + Z_i\hat{z}$. Ils vérifient les relations :

$$(\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \cdot \overrightarrow{V}_1 = (\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \cdot \overrightarrow{V}_2 = 0 \quad ; \quad |\overrightarrow{V}_1|^2 |\overrightarrow{V}_2|^2 = (\overrightarrow{V}_1 \cdot \overrightarrow{V}_2)^2 + |\overrightarrow{V}_1 \wedge \overrightarrow{V}_2|^2 ;$$

$$\overrightarrow{V}_1 \wedge (\overrightarrow{V}_2 \wedge \overrightarrow{V}_3) = (\overrightarrow{V}_1 \cdot \overrightarrow{V}_3)\, \overrightarrow{V}_2 - (\overrightarrow{V}_1 \cdot \overrightarrow{V}_2)\, \overrightarrow{V}_3 \neq (\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \wedge \overrightarrow{V}_3 .$$

La première montre que $\overrightarrow{V}_1 \wedge \overrightarrow{V}_2$ est orthogonal à $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$; sa direction est donnée par la règle du tire-bouchon (*cf.* par exemple $\hat{x} \wedge \hat{y} = \hat{z} = -\hat{y} \wedge \hat{x}$). La seconde permet d'introduire l'angle θ de deux vecteurs (*cf.* section 2.2.1) par

$$\overrightarrow{V}_1 \cdot \overrightarrow{V}_2 = |\overrightarrow{V}_1|\,|\overrightarrow{V}_2| \cos\theta \quad , \quad |\overrightarrow{V}_1 \wedge \overrightarrow{V}_2| = |\overrightarrow{V}_1|\,|\overrightarrow{V}_2| \sin\theta ,$$

et donc de traduire leur **orthogonalité** et leur **parallélisme** respectivement par $\overrightarrow{V}_1 \cdot \overrightarrow{V}_2 = 0$ et $\overrightarrow{V}_1 \wedge \overrightarrow{V}_2 = 0$; on en déduit aussi que lorsque les $\overrightarrow{V}_i$ sont des vecteurs de l'espace ordinaire, $|\overrightarrow{V}_1 \wedge \overrightarrow{V}_2|$ est égal à l'**aire du parallélogramme** construit sur $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$, et que $|(\overrightarrow{V}_1, \overrightarrow{V}_2, \overrightarrow{V}_3)| = |(\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \cdot \overrightarrow{V}_3|$ est égal au **volume du parallélépipède** construit sur $\overrightarrow{V}_1$, $\overrightarrow{V}_2$ et $\overrightarrow{V}_3$, produit de la "surface de base" $|\overrightarrow{V}_1 \wedge \overrightarrow{V}_2|$ par le module de la composante de $\overrightarrow{V}_3$ perpendiculaire à cette base (hauteur du parallélépipède) (figure 14). La dernière montre que le produit vectoriel n'est pas associatif, d'où l'importance des parenthèses et de leurs places ; $\overrightarrow{V}_1 \wedge (\overrightarrow{V}_2 \wedge \overrightarrow{V}_3)$, orthogonal à $\overrightarrow{V}_2 \wedge \overrightarrow{V}_3$, appartient au "plan" $(\overrightarrow{V}_2, \overrightarrow{V}_3)$ (figure 15).

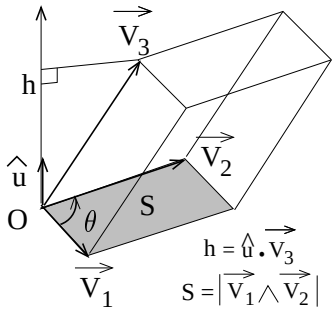


FIGURE 14

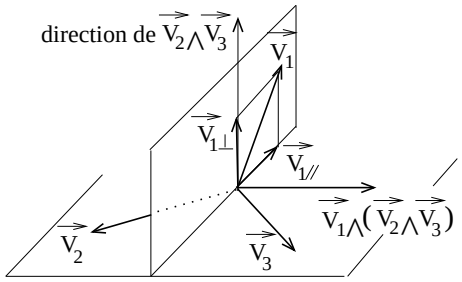
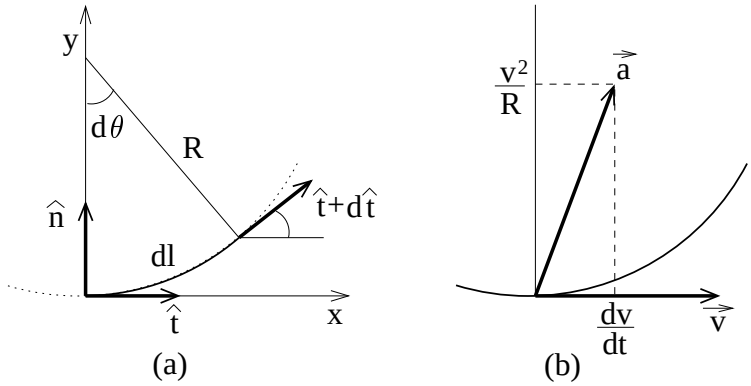


FIGURE 15

Il résulte de la multilinéarité de ces produits que leurs règles de dérivation (lorsque les vecteurs dépendent d'un paramètre) ou de différentiation sont analogues à celles des

produits de fonctions ; par exemple :

$$\begin{aligned}\mathrm{d}(\overrightarrow{V}_1,\overrightarrow{V}_2,\overrightarrow{V}_3) &= \mathrm{d}\big((\overrightarrow{V}_1\wedge\overrightarrow{V}_2)\cdot\overrightarrow{V}_3\big)\\ &= \big(\mathrm{d}\overrightarrow{V}_1\wedge\overrightarrow{V}_2+\overrightarrow{V}_1\wedge\mathrm{d}\overrightarrow{V}_2\big)\cdot\overrightarrow{V}_3+(\overrightarrow{V}_1\wedge\overrightarrow{V}_2)\cdot\mathrm{d}\overrightarrow{V}_3\\ &= (\mathrm{d}\overrightarrow{V}_1,\overrightarrow{V}_2,\overrightarrow{V}_3)+(\overrightarrow{V}_1,\mathrm{d}\overrightarrow{V}_2,\overrightarrow{V}_3)+(\overrightarrow{V}_1,\overrightarrow{V}_2,\mathrm{d}\overrightarrow{V}_3).\end{aligned}$$



FIGURES 16

Une conséquence importante est que la différentielle d'un **vecteur unitaire** (ou de norme constante) est orthogonale à ce vecteur : $\hat{u}\cdot\hat{u}=1$ entraîne $\mathrm{d}(\hat{u}\,.\,\hat{u})=\mathrm{d}\hat{u}\cdot\hat{u}+\hat{u}\cdot\mathrm{d}\hat{u}=2\hat{u}\cdot\mathrm{d}\hat{u}=0$. Si $\hat{t}$ est le vecteur unitaire tangent à une courbe (figure 16a), la relation $\dfrac{\mathrm{d}\hat{t}}{\mathrm{d}l}=\dfrac{1}{R}\,\hat{n}$ définit le vecteur normal et le **rayon de courbure**. L'élément $\mathrm{d}l$ est alors assimilable à un arc de cercle (**cercle osculateur**) et $\dfrac{\mathrm{d}l}{R}$ est l'angle $\mathrm{d}\theta$ dont a tourné $\hat{t}$. On en déduit, en mécanique (classique ou relativiste), la décomposition en composantes tangentielle et normale de l'accélération et de la force (figure 16b) :

$$\vec{a}=\frac{\mathrm{d}\vec{v}}{\mathrm{d}t}=\frac{\mathrm{d}v}{\mathrm{d}t}\,\hat{t}+\frac{v^2}{R}\,\hat{n}\quad;\quad\overrightarrow{F}=\frac{\mathrm{d}\vec{p}}{\mathrm{d}t}=\frac{\mathrm{d}p}{\mathrm{d}t}\,\hat{t}+\frac{vp}{R}\,\hat{n}\ .$$

Si au point considéré, pris comme origine 0, les vecteurs $\hat{t}$ et $\hat{n}$ sont selon Ox et Oy, l'équation de la courbe est $y=\frac{x^2}{2R}$ (approximation près de O de $(y-R)^2+x^2=R^2$).

REMARQUE. Vis-à-vis des rotations, qui conservent les longueurs et les angles, le produit scalaire de deux vecteurs est un invariant ($\mathrm{R}\overrightarrow{V}_1\cdot\mathrm{R}\overrightarrow{V}_2=\overrightarrow{V}_1\cdot\overrightarrow{V}_2$), tandis que le produit vectoriel se comporte comme un vecteur ($\mathrm{R}\overrightarrow{V}_1\wedge\mathrm{R}\overrightarrow{V}_2=\mathrm{R}(\overrightarrow{V}_1\wedge\overrightarrow{V}_2)$). On démontre que tout invariant (aussi appelé "scalaire"), construit à partir de deux vecteurs $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$ est une fonction de $\overrightarrow{V}_1^2$, $\overrightarrow{V}_2^2$ et $\overrightarrow{V}_1\cdot\overrightarrow{V}_2$.

▷ Pseudovecteurs et symétries discrètes

Dans une symétrie S_O, des vecteurs translation $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$ sont changés en $-\overrightarrow{V}_1$ et $-\overrightarrow{V}_2$; cette loi de transformation caractérise les "vrais" vecteurs, appelés aussi **vecteurs polaires**. Par contre $\overrightarrow{V}=\overrightarrow{V}_1\wedge\overrightarrow{V}_2$ reste inchangé ; on l'appelle pseudovecteur ou **vecteur**

axial, et on l'écrit parfois avec une flèche incurvée pour le distinguer des "vrais" vecteurs. Le produit vectoriel d'un vecteur avec un pseudovecteur est un vecteur, tandis que le produit vectoriel de deux pseudovecteurs est un pseudovecteur. On fera attention à l'action "surprenante" d'une symétrie plane $S_P = S_\Delta S_O$ sur un pseudovecteur : sa composante perpendiculaire au plan reste inchangée tandis que sa composante parallèle au plan change de signe ; les figures 17a et 17b représentent l'action de S_0 et de S_P respectivement sur un vrai vecteur et sur un pseudovecteur.

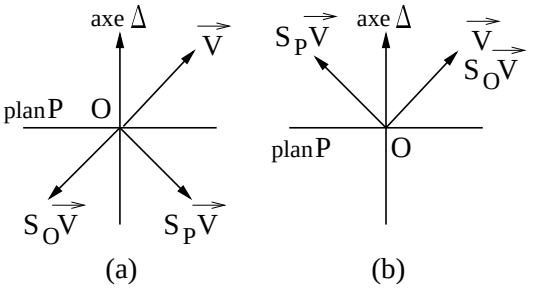


FIGURE 17

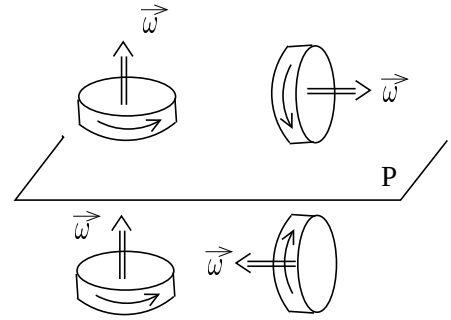


FIGURE 18

Des exemples de pseudovecteurs sont : $\overrightarrow{L} = \overrightarrow{r} \wedge \overrightarrow{p}$ (moment cinétique), $\overrightarrow{\Gamma} = \overrightarrow{r} \wedge \overrightarrow{F}$ (moment d'une force), $\overrightarrow{\omega}$ (vitesse de rotation, figure 18), $\overrightarrow{\mu}$ (moment magnétique), $\overrightarrow{B}$ (champ magnétique), *etc.*

La figure 19 illustre l'**invariance de l'électromagnétisme** dans une symétrie S_P : $\overrightarrow{B'}(S_P\,\overrightarrow{r}) = S_P\,\overrightarrow{B}(\overrightarrow{r})$. La figure 20a illustre la **non invariance des interactions faibles** (désintégration β) : une source de cobalt polarisée perpendiculairement au plan P, donc caractérisée par $\overrightarrow{\mu} \perp P$ laissé invariant par S_P, émet des nombres différents d'électrons de part et d'autre de P ; la désintégration symétrique qui correspond à la figure 20b n'a jamais été observée.

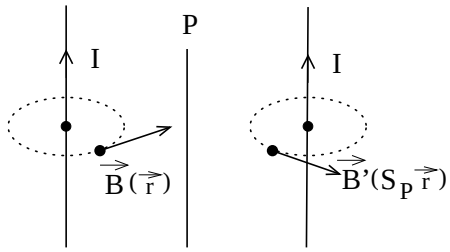


FIGURE 19

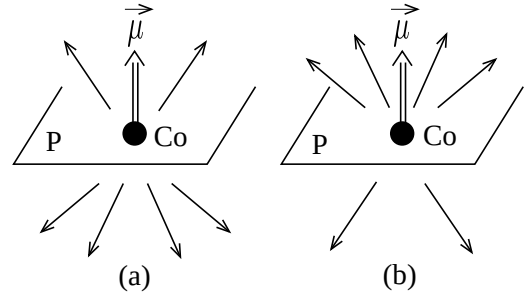


FIGURE 20

REMARQUE. Le produit scalaire d'un vecteur et d'un pseudovecteur, qui change de signe dans une symétrie S_O, est appelé **pseudoscalaire**.

3.2.2 Equations de plans ; fréquences spatiales ; réseaux

▷ Equations de plans ; ondes planes

Un plan orthogonal au vecteur unitaire $\hat{n}(\alpha, \beta, \gamma)$ et distant de d de l'origine O a pour équation (figure 21) :

$$d = \overline{OH} = \hat{n} \cdot \overrightarrow{r} = \alpha x + \beta y + \gamma z \ .$$

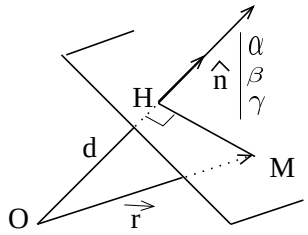


FIGURE 21

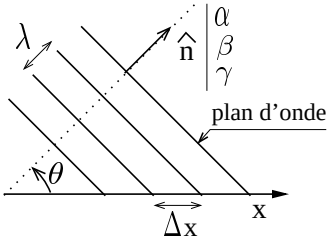


FIGURE 22

Par exemple (figure 22) une onde $a\,\cos(\omega t - \overrightarrow{k} \cdot \overrightarrow{r} + \varphi)$ est dite onde plane parcequ'elle prend une même valeur lorsque $\omega t - \overrightarrow{k} \cdot \overrightarrow{r} + \varphi = \Phi$ (fixé modulo 2π), c'est-à-dire sur les plans d'onde orthogonaux à $\hat{n} = \dfrac{\overrightarrow{k}}{|\overrightarrow{k}|}$ et d'équations $\hat{n} \cdot \overrightarrow{r} = \dfrac{\omega t + \varphi - \Phi}{|\overrightarrow{k}|}$ (modulo $\lambda = \dfrac{2\pi}{|\overrightarrow{k}|}$). A un instant donné ces plans sont équidistants de la longueur d'onde λ ; ils se déplacent à la vitesse de phase $v_\varphi = \dfrac{\omega}{|\overrightarrow{k}|}$. Le vecteur $\overrightarrow{\sigma} = \dfrac{\hat{n}}{\lambda} = \dfrac{\overrightarrow{k}}{2\pi}$ est appelé **fréquence spatiale**. Sa composante $\sigma_x = \dfrac{\alpha}{\lambda} = \dfrac{k_x}{2\pi}$ est bien l'inverse de la période $\Delta x = \dfrac{2\pi}{k_x}$ mesurée sur l'axe Ox ; par contre $\hat{n}\lambda$ ne peut pas représenter un "vecteur période spatiale" car $\alpha\lambda \neq \Delta x$ (sauf à une dimension).

De même l'**interférence de deux ondes planes** $a_i\,\cos(\omega t - \overrightarrow{k}_i \cdot \overrightarrow{r} + \varphi_i)$ conduit à des surfaces d'égale intensité $(\overrightarrow{k}_2 - \overrightarrow{k}_1) \cdot \overrightarrow{r} =$ constant (modulo 2π). Ce sont des plans fixes orthogonaux à $\overrightarrow{k}_2 - \overrightarrow{k}_1$ et distants de $\dfrac{2\pi}{|\overrightarrow{k}_2 - \overrightarrow{k}_1|} \sim \dfrac{\lambda}{\theta}$ (si l'angle θ entre $\overrightarrow{k}_2$ et $\overrightarrow{k}_1$ est petit) (figure 23).

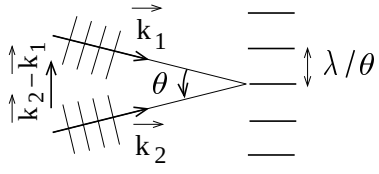


FIGURE 23

▷ Réseaux

Un autre exemple est celui des **plans réticulaires** d'un réseau. On appelle réseau un ensemble de points déduits de O par les translations de vecteurs

$$\overrightarrow{t}_{(m,n,p)} = m\overrightarrow{a} + n\overrightarrow{b} + p\overrightarrow{c} \qquad (m,n,p \in \mathbb{Z})$$

($\overrightarrow{a}$, $\overrightarrow{b}$, $\overrightarrow{c}$ vecteurs de base du réseau). Une famille de plans réticulaires permet par définition de regrouper les points du réseau par plans ; la figure 24a représente une coupe de familles de plans réticulaires, en supposant $\overrightarrow{c}$ perpendiculaire à $\overrightarrow{a}$ et $\overrightarrow{b}$.

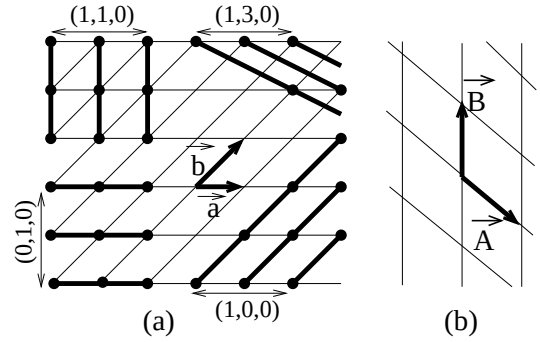


FIGURE 24

Pour décrire ces familles de plans, on introduit le **réseau réciproque** des fréquences spatiales

$$\overrightarrow{\sigma}_{(h,k,l)} = h\overrightarrow{A} + k\overrightarrow{B} + l\overrightarrow{C} \qquad (h,k,l \in \mathbb{Z}) \ ,$$

dont les vecteurs de base $\overrightarrow{A}$, $\overrightarrow{B}$ et $\overrightarrow{C}$ sont définis par $\overrightarrow{A}\cdot\overrightarrow{a} = \overrightarrow{B}\cdot\overrightarrow{b} = \overrightarrow{C}\cdot\overrightarrow{c} = 1$ et $\overrightarrow{A}\cdot\overrightarrow{b} = \overrightarrow{A}\cdot\overrightarrow{c} = 0$, $\overrightarrow{B}\cdot\overrightarrow{c} = \overrightarrow{B}\cdot\overrightarrow{a} = 0$ et $\overrightarrow{C}\cdot\overrightarrow{a} = \overrightarrow{C}\cdot\overrightarrow{b} = 0$ (figure 24b). L'équation d'une famille de plans réticulaires s'écrit alors :

$$\overrightarrow{\sigma}_{(h,k,l)}\cdot\overrightarrow{r} = \mu \qquad \mu \in \mathbb{Z} \ ;$$

elle est caractérisée par les trois nombres h, k et l premiers entre eux appelés indices de Miller (chaque valeur de μ correspondant à un plan de la famille).

DÉMONSTRATION : tout point du réseau appartient bien à la famille car $\overrightarrow{\sigma}_{(h,k,l)}\cdot\overrightarrow{t}_{(m,n,p)} = hm + kn + lp \in \mathbb{Z}$. La condition h, k, l premiers entre eux est nécessaire pour que tous les plans contiennent des points du réseau, notamment le plan $\mu = 1$; en effet le théorème de Bezout assure alors l'existence de nombres entiers m n et p tels $hm+kn+lp = 1$ (contre exemple $2m+2n+4p = 1$ n'a pas de solution). On vérifie aisément sur leurs équations que le plan $\mu = 1$ coupe les trois axes issus de O et parallèles à $\overrightarrow{a}$, $\overrightarrow{b}$ et $\overrightarrow{c}$ en $\frac{\overrightarrow{a}}{h}$, $\frac{\overrightarrow{b}}{k}$ et $\frac{\overrightarrow{c}}{l}$ et que deux plans voisins sont distants de $d_{(h,k,l)} = |\overrightarrow{\sigma}_{(h,k,l)}|^{-1}$.

3.2.3 Différentielles de chemins ; effet Doppler ; lois de Descartes

▷ Différentielle de la longueur AB d'un segment

Si A et B se déplacent de $\mathrm{d}\overrightarrow{r}_A$ et $\mathrm{d}\overrightarrow{r}_B$, la variation de AB est

$$\mathrm{d}(AB) = \mathrm{d}(\hat{u}\cdot\overrightarrow{AB}) = \hat{u}\cdot\mathrm{d}\overrightarrow{AB} = \hat{u}\cdot(\mathrm{d}\overrightarrow{r}_B - \mathrm{d}\overrightarrow{r}_A) \qquad \Big(\hat{u} = \frac{\overrightarrow{AB}}{AB}\Big)$$

(car $\hat{u}\cdot\mathrm{d}\hat{u} = 0$). On retrouve ce résultat en calculant $\mathrm{d}\sqrt{\overrightarrow{AB}^2} = \dfrac{1}{2AB}\,2\overrightarrow{AB}\cdot\mathrm{d}\overrightarrow{AB}$ ou encore en regardant simplement la figure 25 ($A'B' - AB = \overline{HK} - \overline{AB} = \overline{BK} - \overline{AH}$).

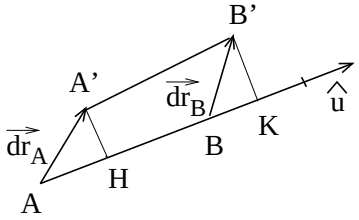


FIGURE 25

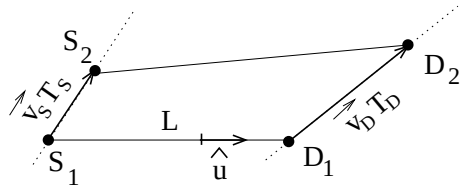


FIGURE 26

EXEMPLE 1. **Effet Doppler** (figure 26). Dans un référentiel $\mathcal{R}$ une source se déplaçant à la vitesse $\overrightarrow{v}_S$ émet des signaux (sonores, optiques...) à intervalles de temps réguliers T_S ; ces signaux sont reçus à intervalles de temps réguliers T_D par un détecteur qui se déplace dans $\mathcal{R}$ à la vitesse $\overrightarrow{v}_D$. Si on suppose $|\overrightarrow{v}_S|T_S$ et $|\overrightarrow{v}_D|T_D \ll L$ (distance de propagation du signal), les vecteurs unitaires $\hat{u}$ des directions de propagation de deux signaux successifs peuvent être confondus. Si V est la vitesse du signal on a $T_D = T_S + \dfrac{\Delta L}{V}$ où $\Delta L = \hat{u}\cdot(\overrightarrow{v}_D T_D - \overrightarrow{v}_S T_S)$ est la différence des trajets parcourus par ces signaux. D'où :

$$T_D = T_S \frac{1 - \frac{\overrightarrow{v}_S\cdot\hat{u}}{V}}{1 - \frac{\overrightarrow{v}_D\cdot\hat{u}}{V}} \,.$$

On remarquera que ce calcul n'implique aucun changement de référentiel. Cependant T_D et T_S ne sont les périodes mesurées dans les référentiels de la source et du détecteur que dans l'approximation galiléenne. En réalité ces périodes (temps propres) sont

$$\tau_{S,D} = T_{S,D}\Big(1 - \frac{v_{S,D}^2}{c^2}\Big)^{\frac{1}{2}}$$

(*cf.* section 4.3.3) et les fréquences correspondantes sont $\nu_{S,D} = \tau_{S,D}^{-1}$.

EXEMPLE 2. **Interférence de deux ondes sphériques** $a_i \cos(\omega t - k r_i + \varphi_i)$ ($r_i = S_i M$ et $k = \dfrac{2\pi}{\lambda}$, figure 27). A grande distance dans une direction proche de celle de S_1S_2, la relation $S_1M - S_2M = S_1S_2 \cos\alpha \sim S_1S_2 \left(1 - \dfrac{\alpha^2}{2}\right)$ entraîne que les lignes d'égale intensité $k(r_2 - r_1) =$ constante (modulo 2π) sont sur des cônes de sommet $S_1 \sim S_2$ (*cf.* aussi section 2.3.2).

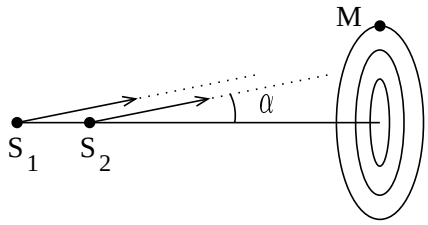


FIGURE 27

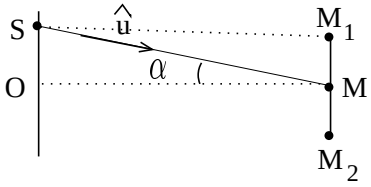


FIGURE 28

EXEMPLE 3. **Largeur de cohérence**. On considère une source de lumière monochromatique étendue centrée en O, dont les points S émettent de façon incohérente, et un point M éloigné (figure 28). On admet que les amplitudes complexes en deux points M_1 et M_2 voisins de M ($\overrightarrow{M_1M_2} \perp \overrightarrow{OM}$) présentent un certain degré de cohérence tant que la distance $SM_1 - SM_2$ varie de moins d'une fraction de longueur d'onde lorsque S décrit la source. Comme $SM_1 - SM_2 = \hat{u} \cdot \overrightarrow{M_1M_2} = \alpha\, M_1M_2$ avec $-\frac{\theta}{2} \leq \alpha \leq \frac{\theta}{2}$ (où θ est l'angle supposé petit sous lequel on voit la source de M), la condition de cohérence correspond approximativement à $M_1M_2 < \frac{\lambda}{\theta}$ (largeur de cohérence).

▷ Différentielles de chemins optiques

Le principe de Fermat pour un dioptre "$n_A\, AM + n_B\, MB$ extremum" (A et B fixes, M point courant sur le dioptre) se traduit par $\mathrm{d}(n_A\, AM + n_B\, MB) = (n_A\, \hat{u}_A - n_B\, \hat{u}_B) \cdot \mathrm{d}\overrightarrow{r}_M = 0$ ($\hat{u}$ vecteur unitaire porté par le rayon) ; $\mathrm{d}\overrightarrow{r}_M$ étant un vecteur tangent au dioptre (figure 29), on obtient la **loi de Descartes** :

$$n_A\, \hat{u}_{A\parallel} = n_B\, \hat{u}_{B\parallel} \qquad (\hat{u}_\parallel \text{ projection de } \hat{u} \text{ sur le plan tangent au dioptre}) \ .$$

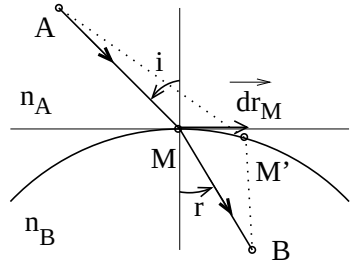


FIGURE 29

Cette loi avec $n_A = n_B$ s'applique aussi à la réflexion sur un miroir. Si maintenant on considère un trajet AMB réellement suivi par la lumière (un rayon) et le chemin

optique $L(A,B) = n_A\,AM + n_B\,MB$ correspondant, sa différentielle s'écrit $\mathrm{d}L(A,B) = n_A\hat{u}_A \cdot (\mathrm{d}\overrightarrow{r}_M - \mathrm{d}\overrightarrow{r}_A) + n_B\hat{u}_B \cdot (\mathrm{d}\overrightarrow{r}_B - \mathrm{d}\overrightarrow{r}_M)$ ou, compte tenu de la loi de Descartes :

$$\mathrm{d}L(A,B) = n_B\,\hat{u}_B \cdot \mathrm{d}\overrightarrow{r}_B - n_A\,\hat{u}_A \cdot \mathrm{d}\overrightarrow{r}_A\ .$$

L'observation de la figure 30 et l'application répétée de cette relation montrent qu'elle s'étend à la traversée d'un nombre quelconque de dioptres ($\mathrm{d}L(A,C) = n_C\,\hat{u}_C \cdot \mathrm{d}\overrightarrow{r}_C - n_A\,\hat{u}_A \cdot \mathrm{d}\overrightarrow{r}_A$), ou même dans la limite continue à un milieu d'indice variable.

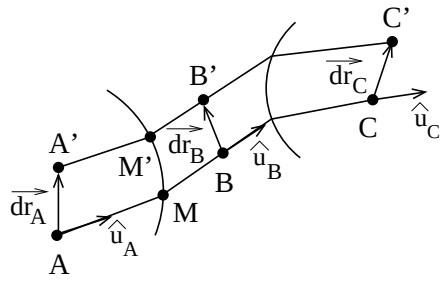


FIGURE 30

On en déduit que tous les points C qui sont situés à une même distance optique d'un point source A sont sur une surface perpendiculaire aux rayons (**Théorème de Malus)**. En effet $\mathrm{d}L = 0$ et $\mathrm{d}\overrightarrow{r}_A = 0$ implique $\hat{u}_C \cdot \mathrm{d}\overrightarrow{r}_C = 0$. (*cf.* aussi section 7.3.2.)

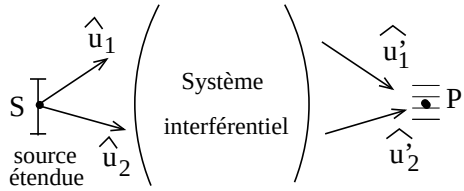


FIGURE 31

Application aux interférences (figure 31). Soient P un point où interfèrent deux ondes issues d'une source ponctuelle S, $\hat{u}_1$, $\hat{u}_2$ les directions d'émission, $\hat{u}'_1$, $\hat{u}'_2$ les directions d'arrivée en P, et L_1 et L_2 les chemins optiques correspondant aux deux rayons allant de S à P. On a :

$$\mathrm{d}(L_2 - L_1) = (\hat{u}'_2 - \hat{u}'_1) \cdot \mathrm{d}\overrightarrow{r}_P - (\hat{u}_2 - \hat{u}_1) \cdot \mathrm{d}\overrightarrow{r}_S\ .$$

Si S est fixé la condition $(\hat{u}'_2 - \hat{u}'_1) \cdot \mathrm{d}\overrightarrow{r}_P =$ constante (modulo λ) donne les surfaces d'égale intensité (déjà obtenues à la section 3.2.2) pour deux ondes planes. On en déduit aussi le lieu de **localisation des interférences** lorsque la source est étendue. En effet pour que l'interférence en P (fixe) ne soit pas détruite lorsqu'on étend la source, il faut que $(\hat{u}_2 - \hat{u}_1) \cdot \mathrm{d}\overrightarrow{r}_S = 0$. Donc soit $\mathrm{d}\overrightarrow{r}_S \perp \hat{u}_2 - \hat{u}_1$ ce qui implique que la source est une fente, soit $\hat{u}_1 = \hat{u}_2$. Dans ce dernier cas, qui suppose un système interférentiel à séparation d'amplitudes, les franges se localisent sur la surface lieu des intersections des rayons issus d'un même rayon incident (exemple figure 12).

3.2.4 Vecteurs surface ; flux de grandeurs

▷ Vecteurs surface

Considérons un parallélogramme orienté construit sur deux vecteurs ordonnés $\overrightarrow{V}_1$ et $\overrightarrow{V}_2$ (figure 32). On appelle "vecteur surface" le vecteur $\overrightarrow{V}_1 \wedge \overrightarrow{V}_2$. La justification de cette appellation est que sa norme est égale à la surface du parallélogramme, et que sa projection $(\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \cdot \hat{u}$ sur une direction $\hat{u}$ correspond à l'**aire algébrique** de la projection du parallélogramme sur un plan perpendiculaire à $\hat{u}$ et orienté par $\hat{u}$ via la règle du tire-bouchon. En effet on vérifie que $(\overrightarrow{V}_1 \wedge \overrightarrow{V}_2) \cdot \hat{u} = (\overrightarrow{V}_{1\parallel} \wedge \overrightarrow{V}_{2\parallel}) \cdot \hat{u}$ où les vecteurs $\overrightarrow{V}_{i\parallel} = \overrightarrow{V}_i - (\overrightarrow{V}_i \cdot \hat{u})\hat{u}$ sont les projections des vecteurs $\overrightarrow{V}_i$ sur le plan.

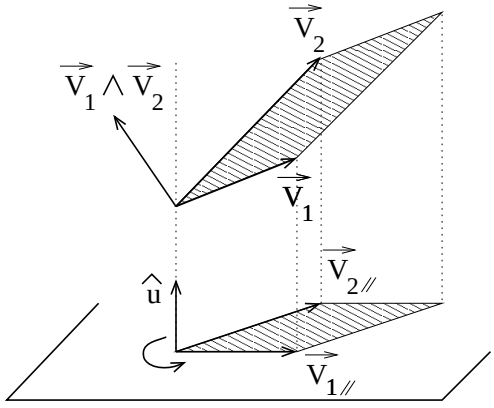


FIGURE 32

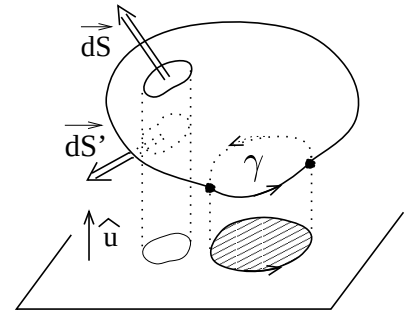


FIGURE 33

Plus généralement pour une surface infinitésimale dS s'appuyant sur une courbe orientée γ, on introduit le vecteur $\mathrm{d}\overrightarrow{S} = \mathrm{d}S\,\hat{n}$ où $\hat{n}$ est le vecteur unitaire normal à l'élément de surface et dont le sens est déterminé par l'orientation de γ ; $\mathrm{d}\overrightarrow{S} \cdot \hat{u}$ est alors l'aire apparente dans la direction $\hat{u}$ ("l'ombre" de dS sur un plan perpendiculaire à $\hat{u}$). Pour une surface finie orientable (*cf.* section 7.2.2), le vecteur surface $\overrightarrow{S}$ est la somme des vecteurs $\mathrm{d}\overrightarrow{S}$; on fera attention à ce que son module $|\overrightarrow{S}|$ ne représente pas l'aire totale, qui est la somme des $|\mathrm{d}\overrightarrow{S}|$. Lorsque la surface s'appuye sur une courbe fermée γ, la somme algébrique des $\mathrm{d}\overrightarrow{S} \cdot \hat{u}$ (aire hachurée sur la figure 33), et donc aussi $\overrightarrow{S}$, ne dépend que de γ. Dans le cas d'une surface fermée, γ est réduit à un point et $\overrightarrow{S} = 0$.

EXEMPLES. $\mathrm{d}\overrightarrow{S} = \dfrac{1}{2}\,\overrightarrow{r} \wedge \mathrm{d}\overrightarrow{r}$ est l'**aire balayée** par un rayon vecteur $\overrightarrow{r}$ (*cf.* section 2.2.1). Le **moment cinétique orbital** $\overrightarrow{L} = \overrightarrow{r} \wedge m\overrightarrow{v}$ s'écrit $2m\dfrac{\mathrm{d}\overrightarrow{S}}{\mathrm{d}t}$. Le **moment magnétique** d'un circuit fermé parcouru par un courant d'intensité I est $\overrightarrow{\mu} = I \oint \dfrac{\overrightarrow{r} \wedge \mathrm{d}\overrightarrow{r}}{2} = I\,\overrightarrow{S}$; cette formule est la version continue de l'expression $\overrightarrow{\mu} = \dfrac{1}{2}\,\Sigma_i\, q_i\,\overrightarrow{r}_i \wedge \overrightarrow{v}_i$ valable pour une distribution discrète de charges ; $\overrightarrow{\mu} = \dfrac{q}{2m}\,\overrightarrow{L}$ pour une seule charge en mouvement orbital.

▷ Flux de particules et flux associés

Considérons des particules d'un type "i" donné, en nombre n_i par unité de volume et ayant une même vitesse $\overrightarrow{v}_i$. Le flux de ces particules à travers $\overrightarrow{\mathrm{d}S}$ est le nombre de particules qui traversent $\overrightarrow{\mathrm{d}S}$ par unité de temps ; c'est le quotient par dt du nombre de particules présentes dans le cylindre de base $\overrightarrow{\mathrm{d}S}$ et de génératrice $\overrightarrow{v}_i\,\mathrm{d}t$, donc de volume $(\overrightarrow{v}_i\,\mathrm{d}t)\cdot\overrightarrow{\mathrm{d}S}$ (figure 34). Si g_i est une grandeur (masse, charge...) attachée aux particules,

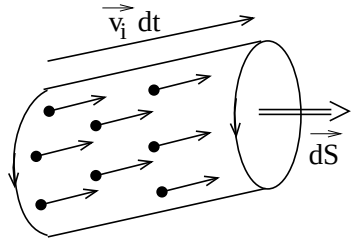


FIGURE 34

le flux de particules s'accompagne d'un flux de cette grandeur égal à $n_i g_i\,(\overrightarrow{v}_i\cdot\overrightarrow{\mathrm{d}S})$. Pour un système macroscopique contenant plusieurs types de particules, ce flux s'obtient en sommant sur "i" et en introduisant des quantités macroscopiques moyennes.

EXEMPLE 1 : le **flux de masse** $\Sigma_i\, n_i m_i\,(\overrightarrow{v}_i\cdot\overrightarrow{\mathrm{d}S})$ s'écrit

$$\mathrm{d}\varphi_m = \rho\,\overrightarrow{v}\cdot\overrightarrow{\mathrm{d}S} = \overrightarrow{j}_m\cdot\overrightarrow{\mathrm{d}S}\ ,$$

où $\rho = \Sigma_i n_i m_i$ est la masse volumique, $\overrightarrow{v} = \dfrac{\Sigma n_i m_i \overrightarrow{v}_i}{\Sigma n_i m_i}$ est la vitesse du centre de masse (vitesse d'ensemble), et $\overrightarrow{j}_m$ est la densité volumique de courant de masse.

EXEMPLE 2 : le **flux de charges électriques** $\Sigma_i\, n_i q\,(\overrightarrow{v}_i\cdot\overrightarrow{\mathrm{d}S})$ (pour une seule espèce de porteurs : $q_i = q$) s'écrit $\rho\,\overrightarrow{v}\cdot\overrightarrow{\mathrm{d}S}$ où $\rho = \Sigma_i n_i q$ est la densité volumique des porteurs et $\overrightarrow{v} = \dfrac{\Sigma n_i \overrightarrow{v}_i}{\Sigma n_i}$ leur vitesse moyenne. Pour plusieurs espèces α de porteurs allant à des vitesses moyennes différentes, il faut ajouter les contributions de chacune :

$$\mathrm{d}\varphi_q = \sum_\alpha \rho_\alpha\,\overrightarrow{v}_\alpha\cdot\overrightarrow{\mathrm{d}S} = \overrightarrow{j}_q\cdot\overrightarrow{\mathrm{d}S} \quad (\overrightarrow{j}_q \text{ densité volumique de courant électrique})\ .$$

EXEMPLE 3 : le **flux de quantité de mouvement** $\Sigma_i\, n_i\,(m_i\overrightarrow{v}_i)\,(\overrightarrow{v}_i\cdot\overrightarrow{\mathrm{d}S})$ s'écrit

$$\mathrm{d}\overrightarrow{\varphi}_p = \rho\overrightarrow{v}\,(\overrightarrow{v}\cdot\overrightarrow{\mathrm{d}S}) + P\,\overrightarrow{\mathrm{d}S}$$

si la distribution des vitesses relatives $\overrightarrow{V}_i = \vec{v}_i - \vec{v}$ est isotrope ($\vec{v}$ étant la vitesse d'ensemble). Ce flux apparaît comme la somme d'un flux macroscopique $\rho\vec{v}\,(\vec{v}\cdot\overrightarrow{\mathrm{d}S})$ associé au mouvement d'ensemble des particules et d'un flux "microscopique"

$\Sigma_i\, n_i\, m_i \overrightarrow{V}_i\,(\overrightarrow{V}_i \cdot \overrightarrow{\mathrm{d}S}) = P\,\overrightarrow{\mathrm{d}S}$ associé à leur mouvement relatif. (Démonstration : on développe $\sum_i n_i m_i(\vec{v}+\overrightarrow{V}_i)\big((\vec{v}+\overrightarrow{V}_i)\cdot\overrightarrow{\mathrm{d}S}\big)$ en utilisant les égalités $\sum_i n_i m_i \overrightarrow{V}_i = 0$, $\sum_i n_i m_i(V_{ix}V_{iy}$ ou $V_{ix}V_{iz}$ ou $V_{iy}V_{iz}) = 0$ et $\displaystyle\sum_i n_i m_i(V_{ix}^2$ ou V_{iy}^2 ou $V_{iz}^2) = \frac{1}{3}\sum_i n_i m_i \overrightarrow{V}_i^{\,2}$ conséquences de l'isotropie.) Pour des particules identiques on obtient l'expression bien connue $P = \dfrac{1}{3}\,nm < \overrightarrow{V}_i^{\,2} >$ de la **pression cinétique** d'un gaz parfait. Dans le cas d'un gaz de photons isotrope, le flux s'écrit $\Sigma_i n_i\Big(\dfrac{e_i}{c}\hat{u}_i\Big)(c\hat{u}_i\cdot\overrightarrow{\mathrm{d}S})$, d'où la relation $P = \dfrac{1}{3}\,u$ entre la **pression de radiation** et la densité volumique d'énergie $u = \Sigma_i\, n_i e_i$.

EXEMPLE 4 : **flux d'énergie**. Si $e_i = \frac{1}{2}m_i v_i^2 + m_i g z_i + \epsilon_i$ est l'énergie totale d'une particule, un petit calcul montre que le flux d'énergie $\Sigma_i n_i e_i\,(\overrightarrow{v}_i\cdot\overrightarrow{\mathrm{d}S})$ associé, dans le cas où la distribution des vitesses relatives $\overrightarrow{V}_i = \overrightarrow{v}_i - \overrightarrow{v}$ est isotrope, s'écrit

$$\mathrm{d}\varphi_E = \rho\,(\tfrac{1}{2}v^2 + e + gz)\,\overrightarrow{v}\cdot\overrightarrow{\mathrm{d}S} + \overrightarrow{v}\cdot P\overrightarrow{\mathrm{d}S}\ ,$$

où $e = \rho^{-1}\Sigma_i n_i\big(\frac{1}{2}m_i V_i^2 + \epsilon_i\big)$ est l'énergie interne massique. Pour un écoulement stationnaire, le flux de masse $\rho\,\overrightarrow{v}\cdot\overrightarrow{\mathrm{d}S}$ étant le même le long d'un tube de fluide, on obtient (si le fluide est isolé) le **théorème de Bernouilli** : la quantité $\frac{1}{2}v^2 + e + gz + \frac{P}{\rho}$ est constante le long d'une ligne de courant. Remarque : dans les exemples 3 et 4 on n'a pas pris en compte les forces à distance entre particules situées de part et d'autre de $\overrightarrow{\mathrm{d}S}$; elles ont pour effet d'ajouter à P une **pression moléculaire**.

▷ Flux d'un champ de vecteurs

Soit $\overrightarrow{A}(\overrightarrow{r})$ la valeur d'un champ au point $\overrightarrow{r}$ où se trouve l'élément de surface $\overrightarrow{\mathrm{d}S}$. On appelle flux de $\overrightarrow{A}$ à travers $\overrightarrow{\mathrm{d}S}$ la quantité $\overrightarrow{A}(\overrightarrow{r})\cdot\overrightarrow{\mathrm{d}S}$. On notera que les flux de masse, charge, énergie et quantité de mouvement (à condition de procéder composante par composante), se mettent bien sous cette forme. On peut aussi citer le flux thermique d'énergie $\overrightarrow{j}_{th}\cdot\overrightarrow{\mathrm{d}S}$ (*cf.* section 7.2.3) et le flux d'énergie électromagnétique $\mathrm{d}\varphi_{em} = (\overrightarrow{E}\wedge\overrightarrow{H})\cdot\overrightarrow{\mathrm{d}S}$ avec $\overrightarrow{E}\wedge\overrightarrow{H}$ vecteur de Poynting.

3.2.5 Sphère, angle solide et applications

▷ Coordonnées sphériques (figures 35 et 36)

Ce sont les coordonnées r, θ et φ définies par

$$x = r\,\sin\theta\,\cos\varphi \quad , \quad y = r\,\sin\theta\,\sin\varphi \quad , \quad z = r\,\cos\theta\ .$$

$\theta \in [0,\pi]$ est la **colatitude** (la **latitude** est $\dfrac{\pi}{2} - \theta$) et $\varphi \in [0, 2\pi[$ la **longitude** (ou angle azimuthal). Pour la sphère $x^2 + y^2 + z^2 = r^2$, le point $\theta = 0$ est le **pôle nord**, $\theta = \pi$ le **pôle sud**, et $\theta = \dfrac{\pi}{2}$ correspond à l'**équateur**. Quand r, θ et φ varient de façon infinitésimale, le point M se déplace respectivement de $\mathrm{d}r$, $r\,\mathrm{d}\theta$ et $r\,\sin\theta\,\mathrm{d}\varphi$ dans les directions définies par les vecteurs unitaires $\hat{r}$ ("verticale"), $\hat{\theta}$ ("sud") et $\hat{\varphi}$ ("est") formant un trièdre direct. On en déduit que des éléments de longueur $\mathrm{d}l$, de surface $\mathrm{d}S$ sur la sphère, et de volume $\mathrm{d}V$ s'écrivent :

$$\mathrm{d}l = \sqrt{\mathrm{d}r^2 + r^2\,(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)} \quad , \quad \mathrm{d}S = r^2\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \quad , \quad \mathrm{d}V = r^2\,\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi\ .$$

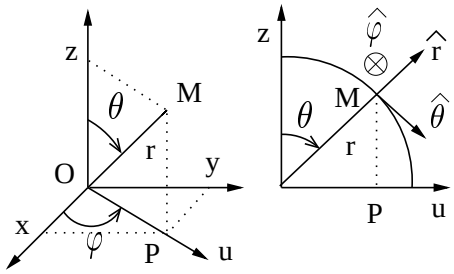


FIGURE 35

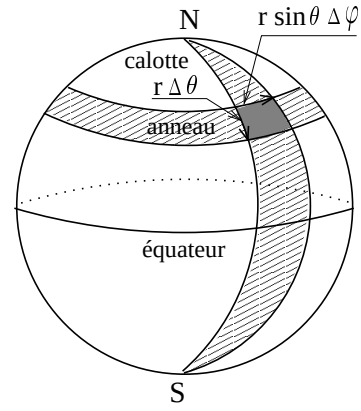


FIGURE 36

EXEMPLES DE LONGUEURS : $r\,\Delta\theta$ pour un arc de **méridien** (φ fixé) et $r\,\sin\theta\,\Delta\varphi$ pour un arc de **parallèle** (θ fixé). La forme quadratique $r^2\,(\mathrm{d}\theta^2+\sin^2\theta\,\mathrm{d}\varphi^2)$ est appelée **métrique** de la sphère de rayon r.

EXEMPLES DE SURFACES : $2\pi r^2\,\sin\theta\,\mathrm{d}\theta$ pour un **anneau** infinitésimal et donc $2\pi r^2\,|\cos\theta_2-\cos\theta_1|$ pour un anneau ($\theta\in[\theta_1,\theta_2]$) ; $2\pi r^2(1-\cos\theta_0)$ pour un **calotte** ($\theta\in[0,\theta_0]$, surface découpée sur la sphère par le cône d'angle au sommet θ_0) ; $2r^2\,\Delta\varphi$ pour un **secteur** ("quartier d'orange" $\varphi\in[\varphi_1,\varphi_2]$) ; $2\pi r^2$ pour une demi sphère (secteur avec $\Delta\varphi=\pi$ ou calotte avec $\theta_0=\dfrac{\pi}{2}$) et $4\pi r^2$ pour la sphère.

EXEMPLES DE VOLUMES : les volumes des cônes de sommet O et limités par une surface sphérique quelconque d'aire S s'obtiennent en remplaçant r^2 par $\dfrac{r^3}{3}$ dans S (intégration de $r^2\,\mathrm{d}r$) ; par exemple $V=\dfrac{4}{3}\pi r^3$ pour la sphère.

▷ Angle solide

Toutes les surfaces précédentes divisées par r^2 définissent l'angle solide $\Omega=\dfrac{S}{r^2}$ du cône de sommet O qui "découpe" la surface S sur la sphère. Plus généralement, pour un cône s'appuyant sur une surface orientée $\overrightarrow{\mathrm{d}S}$ (figure 37), on définit un angle solide élémentaire algébrique :

$$\mathrm{d}\Omega=\frac{\overrightarrow{\mathrm{d}S}\cdot\hat{r}}{r^2}\ .$$

$|\overrightarrow{\mathrm{d}S}\cdot\hat{r}|$ est la surface découpée sur la sphère de rayon r. Le volume de ce cône est $\dfrac{1}{3}\,r^3\,\mathrm{d}\Omega=\dfrac{1}{3}\,r\,\mathrm{d}S\,\cos\alpha=\dfrac{1}{3}\,h\,\mathrm{d}S$.

Applications. Le volume d'un cône quelconque de hauteur h et de base plane d'aire S vaut $\dfrac{1}{3}\,Sh$. Le flux à travers une surface $\overrightarrow{\mathrm{d}S}$ du champ électrique dû à une charge placée à l'origine est $\dfrac{q}{4\pi\epsilon_0 r^2}\,\hat{r}\cdot\overrightarrow{\mathrm{d}S}=\dfrac{q}{4\pi\epsilon_0}\,\mathrm{d}\Omega$. Donc pour une surface fermée orientée vers l'extérieur et entourant une charge totale Q, le flux total vaut $\dfrac{Q}{\epsilon_0}$: **théorème de**

Gauss. Les charges extérieures ne contribuent pas car pour elles l'addition algébrique des angles solides donne zéro (figure 38).

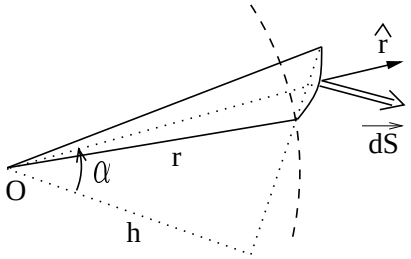


FIGURE 37

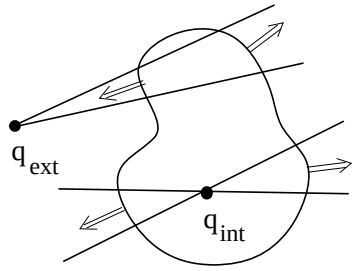


FIGURE 38

▷ **Etendue d'un pinceau lumineux** (figure 39)

Considérons un pinceau, ensemble de rayons rectilignes, traversant une surface $\overrightarrow{\mathrm{d}S}$ et présentant en tout point O de $\overrightarrow{\mathrm{d}S}$ une dispersion angulaire $\mathrm{d}\Omega$ autour de la direction moyenne $\hat{u}$; son étendue est :

$$\mathrm{d}e = (\overrightarrow{\mathrm{d}S} \cdot \hat{u})\,\mathrm{d}\Omega \ .$$

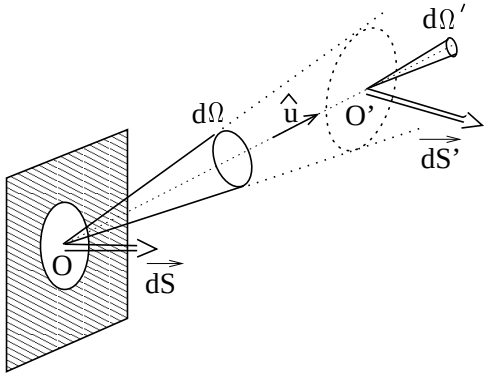


FIGURE 39

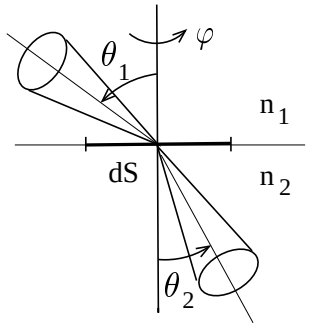


FIGURE 40

La propriété importante de l'étendue est d'être indépendante de la section considérée. En effet à la distance r de $\overrightarrow{\mathrm{d}S}$ la section apparente $\overrightarrow{\mathrm{d}S'} \cdot \hat{u} = r^2\,\mathrm{d}\Omega$ est différente, mais cette différence est compensée par le fait que en tout point O' de $\overrightarrow{\mathrm{d}S'}$ la dispersion des rayons issus de $\overrightarrow{\mathrm{d}S}$ est devenue $\mathrm{d}\Omega' = \frac{\overrightarrow{\mathrm{d}S}\cdot\hat{u}}{r^2}$. Quand un pinceau traverse un dioptre (figure 40), la quantité $n^2\,(\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi)\,(\mathrm{d}S\cos\theta)$ reste constante à cause des lois de Descartes ($n_1\,\sin\theta_1 = n_2\,\sin\theta_2$ et $\varphi_1 = \varphi_2$) et de leurs différentielles ($n_1\,\cos\theta_1\,\mathrm{d}\theta_1 = n_2\,\cos\theta_2\,\mathrm{d}\theta_2$ et $\mathrm{d}\varphi_1 = \mathrm{d}\varphi_2$). Donc l'**étendue optique** n^2 de du pinceau se conserve. Il en est de même de sa **luminance** définie par $L = \frac{\mathrm{d}F}{n^2\,\mathrm{d}e}$ où $\mathrm{d}F$ est le flux d'énergie, supposé constant en négligeant les pertes en cours de propagation. Une conséquence non triviale est que l'énergie et donc le **nombre de photons** par cellule élémentaire de l'espace de phase, émis par une source (thermique, laser, *etc.*), se conserve lui aussi dans la propagation : c'est une caractéristique de la source.

DÉMONSTRATION : considérons un pinceau quasi-monochromatique de luminance $\mathrm{d}L = l(\omega)\,\mathrm{d}\omega$ ($l(\omega)$ luminance spectrale) et de section droite $\mathrm{d}x\,\mathrm{d}y$ qui se propage dans le vide selon Oz ($\omega = ck_z$). Soient $\mathrm{d}\omega = c\,\mathrm{d}k_z$ et $\mathrm{d}k_x$, $\mathrm{d}k_y$, $\mathrm{d}k_z$ les dispersions en pulsation et vecteur d'onde des photons associés ; alors

$\mathrm{d}F = l(\omega)\,\mathrm{d}\omega\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\Omega$ avec $\mathrm{d}\Omega = \frac{\mathrm{d}k_x\,\mathrm{d}k_y}{k_z^2}$. Le nombre de photons contenus dans un volume $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ (avec $\mathrm{d}z = c\,\mathrm{d}t$) est : $\mathrm{d}N = (\hbar\omega)^{-1}\,\mathrm{d}F\,\mathrm{d}t = \frac{c^2}{\hbar\omega^3}\,l(\omega)\,\mathrm{d}k_x\,\mathrm{d}k_y\,\mathrm{d}k_z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. Comme une cellule correspond à $\mathrm{d}k_x\,\mathrm{d}k_y\,\mathrm{d}k_z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = (2\pi)^3$ (équivalent à $\mathrm{d}^3r\,\mathrm{d}^3p = h^3$), ce nombre par cellule est $(2\pi)^3\frac{c^2}{\hbar\omega^3}\,l(\omega)$; il se conserve (comme ω et $l(\omega)$).

REMARQUE : pour une source thermique, ce nombre n'est autre que le facteur de Planck $\left(\exp\frac{\hbar\omega}{k_BT}-1\right)^{-1}$ multiplié par deux (pour tenir compte des deux états d'hélicité des photons ; *cf.* section 4.4.1).

3.2.6 Géométrie sphérique ; notion de transport parallèle

▷ Géodésiques

Sur la sphère les géodésiques (plus court chemin d'un point à un autre) sont les grands cercles intersections avec les plans passant par O, les méridiens par exemple mais pas les parallèles à l'exception de l'équateur. Les **triangles sphériques**, triangles dont les côtés (de longueurs a, b, c) sont des arcs de grands cercles, ont des propriétés trigonométriques qui ressemblent à celles des triangles plans (dont les côtés rectilignes sont des géodésiques du plan) ; par exemple :

$$\frac{\sin\widehat{A}}{\sin(a/R)} = \frac{\sin\widehat{B}}{\sin(b/R)} = \frac{\sin\widehat{C}}{\sin(c/R)} \quad ; \quad \cos\tfrac{c}{R} = \cos\tfrac{a}{R}\,\cos\tfrac{b}{R} + \sin\tfrac{a}{R}\,\sin\tfrac{b}{R}\,\cos\widehat{A}\ .$$

Pour $R \gg a, b, c$ ces formules donnent celles bien connues du plan $\frac{\sin\widehat{A}}{a} = \frac{\sin\widehat{B}}{b} = \frac{\sin\widehat{C}}{c}$ et $c^2 = a^2 + b^2 - 2ab\cos\widehat{A}$. Une propriété facile à vérifier est que la somme des angles d'un triangle sphérique vaut :

$$\widehat{A} + \widehat{B} + \widehat{C} = \sigma = \pi + \tfrac{S}{R^2} \quad (S \text{ aire du triangle } ABC)\ .$$

On remarque en effet en coupant un triangle en deux que $\sigma = \sigma_1 + \sigma_2 - \pi$ (figure 41a) ce qui suggère, puisque $\sigma - \pi = (\sigma_1 - \pi) + (\sigma_2 - \pi)$, que $\sigma - \pi$ est proportionnel à l'aire du triangle et même plus précisément à $\frac{S}{R^2}$ puisque σ est sans dimension ; reste à vérifier que le coefficient de proportionnalité vaut 1 ce qui est évident pour un triangle trirectangle : sa surface vaut le huitième de celle de la sphère (figure 41b) et on a bien $\sigma = 3\,\frac{\pi}{2} = \pi + \frac{1}{R^2}\,\frac{\pi R^2}{2}$.

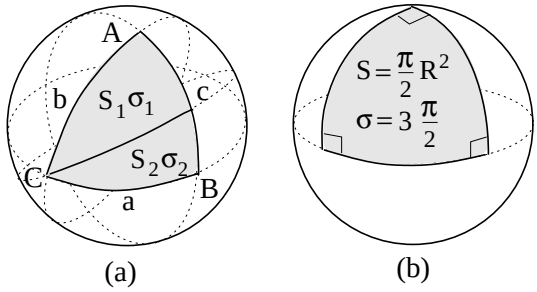


FIGURE 41

▷ Courbure et transport parallèle

De façon générale on définit la **courbure de Gauss** (ou intrinsèque) d'une surface en un point M de la manière suivante. On considère un circuit élémentaire fermé entourant M dont les côtés sont des géodésiques et on transporte le long de ce circuit un vecteur tangent à la surface en maintenant constant la longueur du vecteur et l'angle qu'il fait avec les géodésiques ; on dit alors que le vecteur subit un **transport parallèle** (généralisation naturelle, au cas d'une surface quelconque, du transport d'un vecteur dans le plan euclidien, parallèlement à lui même). Soient $\mathrm{d}S$ la surface délimitée par le circuit et $\mathrm{d}\alpha$ l'angle dont a tourné le vecteur après un tour ; la courbure de Gauss est le rapport $\frac{\mathrm{d}\alpha}{\mathrm{d}S}$ ($\mathrm{d}\alpha$ et $\mathrm{d}S$ sont algébriques). Pour un triangle sphérique $\alpha = \sigma - \pi = \frac{S}{R^2}$ (*cf.* exemple figure 42a) ;

la courbure de la sphère est donc la même en tout point, positive et égale à $\frac{1}{R^2}$. Pour un triangle plan $\alpha = 0$ et la courbure du plan euclidien est évidemment nulle, comme celle du cylindre, ce qui est moins intuitif (figure 42b). La figure 42c montre un exemple de surface à courbure négative.

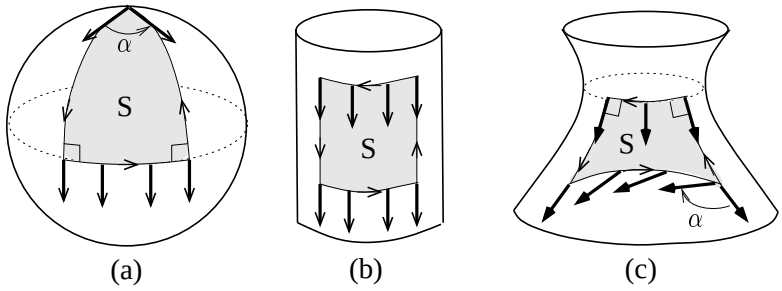


FIGURE 42

REMARQUE. **Transport parallèle et pendule de Foucault**. Le transport parallèle d'un vecteur tangent d'un point A (θ,φ) à un point B voisin $(\theta,\, \varphi + \mathrm{d}\varphi)$ sur le même parallèle ne conserve pas l'angle qu'il fait avec le parallèle car ce dernier n'est pas une géodésique. La variation de cet angle est $-\cos\theta\,\mathrm{d}\varphi$. On obtient ce résultat en considérant le transport le long du chemin géodésique fermé $ACDBA$ qui correspond à une rotation $\delta\alpha = \frac{\text{aire}(ACDBA)}{R^2} = \cos\theta\,\mathrm{d}\varphi$ et en remarquant que le transport de A à B le long de $ACDB$ ne modifie pas l'angle du vecteur avec le parallèle (figure 43a) ; $-\cos\theta\,\mathrm{d}\varphi$ correspond donc bien au seul transport de A à B. Plus intuitivement, en remarquant que l'arc "infinitésimal" de grand cercle de A à B (géodésique) est asssimilable au segment AB et en maintenant l'angle du vecteur avec ce segment constant lors de son transport de A à B, on voit que la variation d'angle avec le parallèle (comme avec le méridien) est (figure 43b) : $-\widehat{APB} = -\frac{AB}{PA} = -\frac{R\,\sin\theta\,\mathrm{d}\varphi}{R\,\mathrm{tg}\theta} = -\cos\theta\,\mathrm{d}\varphi$. Pour un transport fini le long du parallèle, cette variation vaut $-\cos\theta\,\Delta\varphi$.

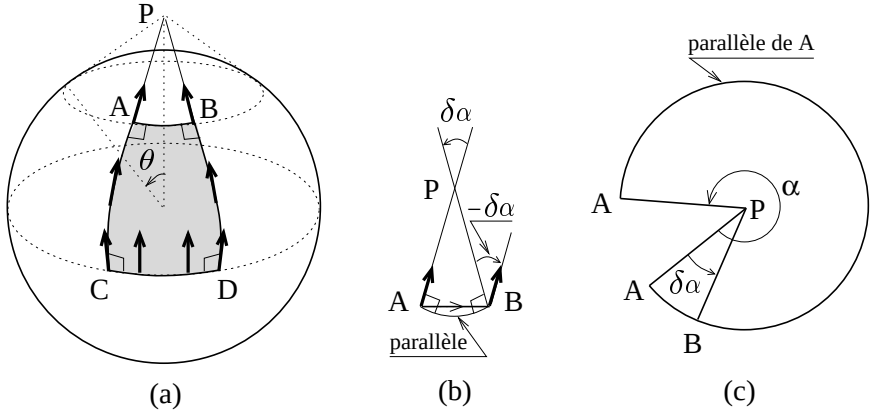


FIGURE 43

Si on admet que la direction d'oscillation d'un pendule de Foucault (tangente au globe terrestre) est transportée parallèlement à elle même lors de la rotation de la terre, $-\cos\theta\,\mathrm{d}\varphi$ est aussi l'angle de rotation du plan d'oscillation du pendule par rapport à la direction nord locale. En une journée, cet angle est $-\alpha = -2\pi\,\cos\theta$ (α angle du cône de sommet P "déployé" ; figure 43c).

3.3 VECTEURS TOURNANTS ; MÉCANIQUE DU SOLIDE

Les rotations et mouvements de rotation jouent un rôle très important en physique (*cf.* aussi section 6.3).

3.3.1 Vecteurs tournants ; changements de référentiels

▷ Représentation vectorielle des rotations

Soit $\hat{n}$ le vecteur unitaire de l'axe de rotation, et soit

$$\overrightarrow{V} = (\overrightarrow{V}\cdot\hat{n})\,\hat{n} + \overrightarrow{V} - (\overrightarrow{V}\cdot\hat{n})\,\hat{n} = \overrightarrow{V}_{\parallel} + \overrightarrow{V}_{\perp}$$

la décomposition d'un vecteur $\overrightarrow{V}$ en vecteurs parallèle et orthogonal à $\hat{n}$ ($\overrightarrow{V}_{\parallel}\wedge\hat{n}=0$; $\overrightarrow{V}_{\perp}\cdot\hat{n}=0$)(figure 44). La rotation laisse $\overrightarrow{V}_{\parallel}=\overrightarrow{OH}$ inchangé et tourne $\overrightarrow{V}_{\perp}=\overrightarrow{HM}$ d'un angle θ dans un plan orthogonal à $\hat{n}$, donc :

$$\overrightarrow{V}' = \mathrm{R}_O(\hat{n},\theta)\,\overrightarrow{V} = \overrightarrow{V}_{\parallel} + \cos\theta\,\overrightarrow{V}_{\perp} + \sin\theta\,\hat{n}\wedge\overrightarrow{V}\ .$$

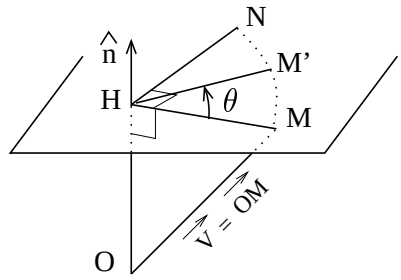


FIGURE 44

($\hat{n}\wedge\overrightarrow{V} = \hat{n}\wedge\overrightarrow{V}_{\perp} = \overrightarrow{HN}$ se déduit de $\overrightarrow{V}_{\perp}$ par une rotation de $\pi/2$). Pour les composantes de $\overrightarrow{V}$, cette égalité correspond à $Z'=Z$, $X'=X\,\cos\theta - Y\,\sin\theta$ et $Y'=X\,\sin\theta + Y\,\cos\theta$, l'axe Z étant choisi selon $\hat{n}$. Pour une **rotation infinitésimale** on a :

$$\mathrm{d}\overrightarrow{V} = \overrightarrow{V}' - \overrightarrow{V} = \mathrm{d}\theta\,\hat{n}\wedge\overrightarrow{V}\ .$$

On en déduit que **l'équation caractéristique d'un vecteur tournant** $\overrightarrow{V}(t)$ est

$$\frac{\mathrm{d}\overrightarrow{V}}{\mathrm{d}t} = \overrightarrow{\omega}\wedge\overrightarrow{V}\ ,$$

où $\overrightarrow{\omega}(t) = \dfrac{\mathrm{d}\theta}{\mathrm{d}t}\,\hat{n}(t)$ est la vitesse angulaire, et que :

$$\frac{\mathrm{d}^2\overrightarrow{V}}{\mathrm{d}t^2} = \frac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t}\wedge\overrightarrow{V} + \overrightarrow{\omega}\wedge\left(\overrightarrow{\omega}\wedge\overrightarrow{V}\right) = \frac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t}\wedge\overrightarrow{V} - \omega^2\overrightarrow{V}_{\perp}\ .$$

Cas particulier : si $\overrightarrow{\omega}$ est constant ($\theta = \omega t$), l'expression de $\overrightarrow{V}(t)$ en fonction de $\overrightarrow{V}_0 = \overrightarrow{V}(t=0)$ est

$$\overrightarrow{V}(t) = \overrightarrow{V}_{0\parallel} + \overrightarrow{V}_{0\perp} \cos \omega t + \hat{n} \wedge \overrightarrow{V}_{0\perp} \sin \omega t \ ,$$

et $\dfrac{\mathrm{d}^2 \overrightarrow{V}}{\mathrm{d}t^2} = -\omega^2 \overrightarrow{V}_{\perp}(t)$ correspond à une **accélération centripète**.

▷ Formules générales des changements de référentiels (figure 45)

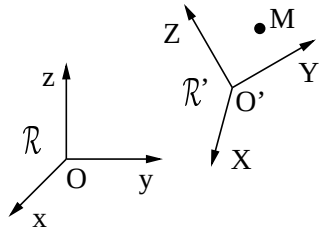


FIGURE 45

Soit $O'XYZ$ un référentiel $\mathcal{R}'$ mobile par rapport au référentiel $\mathcal{R} \equiv Oxyz$, et soit un point M mobile repéré par ses coordonnées $X(t)$, $Y(t)$ et $Z(t)$ dans $\mathcal{R}'$:

$$\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M} = \overrightarrow{OO'} + X\hat{X} + Y\hat{Y} + Z\hat{Z} \ .$$

A l'aide des relations $\dfrac{\mathrm{d}}{\mathrm{d}t} X\hat{X} = X \dfrac{\mathrm{d}\hat{X}}{\mathrm{d}t} + \dot{X}\hat{X}$, $\dfrac{\mathrm{d}^2}{\mathrm{d}t^2} X\hat{X} = X \dfrac{\mathrm{d}^2\hat{X}}{\mathrm{d}t^2} + 2\dot{X} \dfrac{\mathrm{d}\hat{X}}{\mathrm{d}t} + \ddot{X}\hat{X}$et des relations $\dfrac{\mathrm{d}\hat{X}}{\mathrm{d}t} = \overrightarrow{\omega} \wedge \hat{X}$ et $\dfrac{\mathrm{d}^2\hat{X}}{\mathrm{d}t^2} = \dfrac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t} \wedge \hat{X} + \overrightarrow{\omega} \wedge \left(\overrightarrow{\omega} \wedge \hat{X}\right)$ faisant intervenir le vecteur unitaire tournant $\hat{X}$ (et leurs analogues pour Y et Z), on obtient immédiatement :

$$\overrightarrow{v} = \frac{\mathrm{d}\overrightarrow{OM}}{\mathrm{d}t} = \Big[\frac{\mathrm{d}}{\mathrm{d}t}\overrightarrow{OO'} + \overrightarrow{\omega} \wedge \overrightarrow{O'M}\Big] + \big[\dot{X}\hat{X} + \dot{Y}\hat{Y} + \dot{Z}\hat{Z}\big] = \overrightarrow{v}_e + \overrightarrow{v}_r \ ,$$

$$\overrightarrow{a} = \frac{\mathrm{d}^2\overrightarrow{OM}}{\mathrm{d}t^2} = \Big[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\overrightarrow{OO'} + \frac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t} \wedge \overrightarrow{O'M} + \overrightarrow{\omega} \wedge \big(\overrightarrow{\omega} \wedge \overrightarrow{O'M}\big)\Big] + 2\overrightarrow{\omega} \wedge \overrightarrow{v}_r$$
$$+ \big[\ddot{X}\hat{X} + \ddot{Y}\hat{Y} + \ddot{Z}\hat{Z}\big] = \overrightarrow{a}_e + \overrightarrow{a}_c + \overrightarrow{a}_r \ .$$

$\overrightarrow{v}_e$ et $\overrightarrow{a}_e$ sont les **vitesse et accélération d'entraînement**, vitesse et accélération du point M' fixe dans $\mathcal{R}'$ avec lequel M coïncide à l'instant t ; $\overrightarrow{v}_r$ et $\overrightarrow{a}_r$ sont les **vitesse et accélération relatives** ("mesurées" dans $\mathcal{R}'$) ; $\overrightarrow{a}_c = 2\overrightarrow{\omega} \wedge \overrightarrow{v}_r$ est **l'accélération de Coriolis**.

▷ Référentiel local (figure 46)

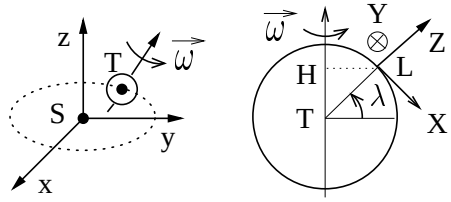


FIGURE 46

Soit $\mathcal{R}$ le référentiel de Copernic (inertiel) centré au centre de masse (c.d.m.) S du système solaire, et $\mathcal{R}'$ le référentiel $LXYZ$ centré en un lieu L de latitude λ à la surface de la Terre et dont les axes sont dirigés vers le sud, l'est et le zénith. $\mathcal{R}'$ tourne à la vitesse $\vec{\omega}$ constante par rapport à $\mathcal{R}$. L'équation du mouvement d'un point matériel M de masse m dans $\mathcal{R}$ s'écrit

$$m(\vec{a}_e + \vec{a}_c + \vec{a}_r) = m(\vec{g}_A(M) + \vec{g}_T(M)) + (\text{autres forces}) \ ,$$

où $\vec{g}_A$ est le champ de gravité dû aux astres du système solaire (Soleil, Lune *etc...*) et $\vec{g}_T$ celui dû à la Terre. Comme le c.d.m. T de la Terre est "en chute libre" autour de S avec l'accélération $\vec{g}_A(T)$ et que $\overrightarrow{TM'}$, fixe dans $\mathcal{R}'$, est un vecteur tournant à la vitesse $\vec{\omega}$ dans $\mathcal{R}$, on a

$$\vec{a}_e = \frac{\mathrm{d}^2\overrightarrow{SM'}}{\mathrm{d}t^2} = \frac{\mathrm{d}^2\overrightarrow{ST}}{\mathrm{d}t^2} + \frac{\mathrm{d}^2\overrightarrow{TM'}}{\mathrm{d}t^2} = \vec{g}_A(T) - \omega^2\,\overrightarrow{HM'} \ ,$$

où H est la projection de M' (ou L) sur l'axe de rotation. Le mouvement de M dans $\mathcal{R}'$ (mouvement relatif) est alors régi par l'équation :

$$\vec{a}_r = \big(\vec{g}_A(M) - \vec{g}_A(T)\big) + \big(\vec{g}_T(M) + \omega^2\,\overrightarrow{HM}\big) - 2\vec{\omega} \wedge \vec{v}_r + \frac{1}{m}\,(\text{autres forces}) \ .$$

Dans le second membre le premier terme est en général négligeable (sauf pour décrire le phénomène de marée), et le second est le **champ de pesanteur** $\vec{g}$ (local). Les effets du terme de Coriolis (**déviation vers l'est et pendule de Foucault**) sont étudiés à la section 6.3.

3.3.2 Référentiel du centre de masse

Les formules cinématiques de changements de référentiels se simplifient grandement lorsque $\mathcal{R}'$ effectue un mouvement de translation ($\vec{\omega} = 0$) par rapport au référentiel $\mathcal{R}$, car alors $\vec{a}_c = 0$, $\vec{v}_e = \vec{v}_{O'}$ et $\vec{a}_e = \vec{a}_{O'}$ ne dépendent pas du point considéré.

▷ Théorèmes de Koenig

En mécanique, pour décrire la dynamique globale d'un système de points matériels M_i de masses m_i, on utilise le référentiel $\mathcal{R}^{cm}$ d'axes parallèles à ceux de $\mathcal{R}$ et d'origine G (c.d.m.) défini par :

$$m\overrightarrow{OG} = \sum_i m_i\overrightarrow{OM_i} \quad (m = \sum_i m_i \text{ masse totale}) \qquad \text{ou} \qquad \sum_i m_i\overrightarrow{GM_i} = 0 \ .$$

$\mathcal{R}^{cm}$ se translate par rapport à $\mathcal{R}$ à la vitesse $\overrightarrow{v}_G$ donnée par $m\overrightarrow{v}_G = \sum_i m_i \overrightarrow{v}_i = \overrightarrow{P}$ (quantité de mouvement totale dans $\mathcal{R}$) ; l'accélération $\overrightarrow{a}_G$ est telle que $m\overrightarrow{a}_G = \sum_i m_i \overrightarrow{a}_i = \sum_i \overrightarrow{f}_i = \overrightarrow{F}$ (force totale dans $\mathcal{R}$). On notera que la quantité de mouvement totale dans $\mathcal{R}^{cm}$ est nulle : $\overrightarrow{P}^{cm} = \sum_i m_i \overrightarrow{v}_i^{cm} = \sum_i m_i(\overrightarrow{v}_i - \overrightarrow{v}_G) = 0$.

Un calcul élémentaire utilisant $\overrightarrow{v}_i^{cm} = \overrightarrow{v}_i - \overrightarrow{v}_G$ montre que le choix de $\mathcal{R}^{cm}$ permet de relier simplement les **énergies cinétiques** $E_c = \sum_i \frac{1}{2} m_i \overrightarrow{v}_i^{\,2}$ et $E_c^{cm} = \sum_i \frac{1}{2} m_i \overrightarrow{v}_i^{cm\,2}$, ainsi que les **moments cinétiques** $\overrightarrow{J}_O = \sum_i \overrightarrow{OM}_i \wedge m_i \overrightarrow{v}_i$ et $\overrightarrow{J}_G^{cm} = \sum_i \overrightarrow{GM}_i \wedge m_i \overrightarrow{v}_i^{cm}$ (aussi égal à $\sum_i \overrightarrow{GM}_i \wedge m_i \overrightarrow{v}_i$) dans $\mathcal{R}$ et $\mathcal{R}^{cm}$:

$$E_c = \frac{1}{2} m \overrightarrow{v}_G^{\,2} + E_c^{cm} \quad ; \quad \overrightarrow{J}_O = \overrightarrow{OG} \wedge m\overrightarrow{v}_G + \overrightarrow{J}_G^{cm} \ .$$

On obtient de même des expressions simples dans $\mathcal{R}^{cm}$ pour les **théorèmes de l'énergie cinétique et du moment cinétique** :

$$\frac{\mathrm{d}E_c^{cm}}{\mathrm{d}t} = \sum_i \overrightarrow{v}_i^{\,cm} \cdot \overrightarrow{f}_i \quad ; \quad \frac{\mathrm{d}\overrightarrow{J}_G^{cm}}{\mathrm{d}t} = \sum_i \overrightarrow{GM}_i \wedge \overrightarrow{f}_i \ .$$

Elles ne font intervenir que les forces $\overrightarrow{f}_i$ dans $\mathcal{R}$, et pas l'accélération d'entraînement, alors que pour chaque masse $m_i \overrightarrow{a}_i^{cm} = \overrightarrow{f}_i - m_i \overrightarrow{a}_G$.

▷ Réduction du problème à deux corps

(système isolé de deux points matériels (M_1, m_1) et (M_2, m_2))

Comme la force totale $\overrightarrow{F} = \overrightarrow{f}_{1\to 2} + \overrightarrow{f}_{2\to 1}$ est nulle, $\mathcal{R}^{cm}$ se déplace à vitesse uniforme. Pour décrire le mouvement dans $\mathcal{R}^{cm}$ on introduit les quantités $\overrightarrow{r} = \overrightarrow{M_1M_2}$, $\mu = \dfrac{m_1 m_2}{m_1 + m_2}$ (**masse réduite**) et $\overrightarrow{f} = \overrightarrow{f}_{1\to 2}$. Alors à l'aide des relations $\overrightarrow{GM_2} = \dfrac{m_1}{m}\overrightarrow{r}$, $\overrightarrow{GM_1} = -\dfrac{m_2}{m}\overrightarrow{r}$, et de leurs dérivées $\overrightarrow{v}_2^{cm} = \dfrac{m_1}{m}\dot{\overrightarrow{r}}$ et $\overrightarrow{v}_1^{cm} = -\dfrac{m_2}{m}\dot{\overrightarrow{r}}$, on obtient facilement :

$$E_c^{cm} = \frac{1}{2}\mu \dot{\overrightarrow{r}}^{\,2} \quad ; \quad \overrightarrow{J}_G^{cm} = \overrightarrow{r} \wedge \mu\dot{\overrightarrow{r}} \quad ; \quad \frac{\mathrm{d}E_c^{cm}}{\mathrm{d}t} = \dot{\overrightarrow{r}} \cdot \overrightarrow{f} \quad ; \quad \frac{\mathrm{d}\overrightarrow{J}_G^{cm}}{\mathrm{d}t} = \overrightarrow{r} \wedge \overrightarrow{f} \ .$$

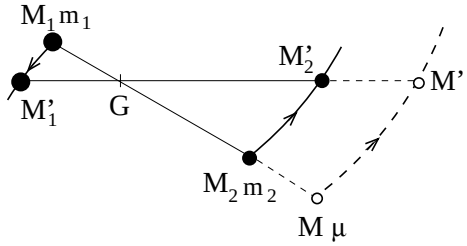


FIGURE 47

La dynamique du vecteur $\overrightarrow{r}$ est donc celle d'un point matériel fictif M, tel que $\overrightarrow{r} = \overrightarrow{GM}$, de masse μ et soumis à la force $\overrightarrow{f}$. On peut aussi obtenir ce résultat en remarquant que

$m_1 \overrightarrow{a}_1 = \overrightarrow{f}_{2\to1}$ et $m_2 \overrightarrow{a}_2 = \overrightarrow{f}_{1\to2}$ entraînent $\mu \ddot{\overrightarrow{r}} = \overrightarrow{f}$ (par exemple $\ddot{\overrightarrow{r}} = -G(m_1 + m_2)\dfrac{\hat{r}}{r^2}$ dans le cas du problème de Kepler ; *cf.* section 2.3.3). Une fois la trajectoire de M connue, celles de M_1 et M_2 dans $\mathcal{R}^{cm}$ s'en déduisent par homothétie (figure 47). En particulier si $\overrightarrow{f}$ est parallèle à $\overrightarrow{r}$, $\overrightarrow{J}^{cm}_G$ est constant (L.C. du moment cinétique). Donc, $\overrightarrow{GM_1}$ et $\overrightarrow{GM_2}$ colinéaires à $\overrightarrow{r}$ étant orthogonaux à $\overrightarrow{J}^{cm}_G$, les trajectoires de M_1 et M_2 sont, dans le cas d'une **force centrale**, situées dans le plan passant par G et perpendiculaire à $\overrightarrow{J}^{cm}_G$.

3.3.3 Cinématique et dynamique d'un corps solide

On se limite à l'introduction des principaux outils nécessaires à la description de la dynamique des solides et à la présentation de deux exemples.

▷ Champ des vitesses d'un solide

Comme les distances entre points d'un corps solide sont par définition invariantes, un vecteur $\overrightarrow{AB}$ joignant deux points du solide ne peut que tourner :

$$\frac{\mathrm{d}\overrightarrow{AB}}{\mathrm{d}t} = \overrightarrow{v}_B - \overrightarrow{v}_A = \overrightarrow{\omega} \wedge \overrightarrow{AB} \ .$$

Comme $\overrightarrow{\omega} \wedge \overrightarrow{AB}$ est orthogonal à $\overrightarrow{\omega}$ il en résulte que, à un instant donné, tous les points du solide ont même composante de vitesse $\overrightarrow{v}_{\parallel}$ parallèle à $\overrightarrow{\omega}$ et que sur un axe parallèle à $\overrightarrow{\omega}$ tous les points ont de plus même composante $\overrightarrow{v}_{\perp}$ orthogonale à $\overrightarrow{\omega}$. En faisant varier B dans la relation $\overrightarrow{v}_{B\perp} = \overrightarrow{v}_{A\perp} + \overrightarrow{\omega} \wedge \overrightarrow{AB}$, on vérifie qu'il existe un point $B \equiv I$ tel que $\overrightarrow{v}_{I\perp} = 0$ (figure 48) ; il y a donc un axe parallèle à $\overrightarrow{\omega}$ et passant par I tel que $\overrightarrow{v}_{\perp} = 0$ pour tous ses points. La relation

$$\overrightarrow{v}_A = \overrightarrow{v}_I + \overrightarrow{\omega} \wedge \overrightarrow{IA}$$

montre, comme déjà établi à la section 3.1.3, que le mouvement d'un solide est une suite de vissages élémentaires.

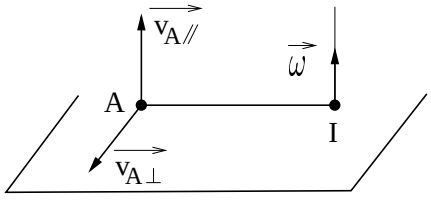


FIGURE 48

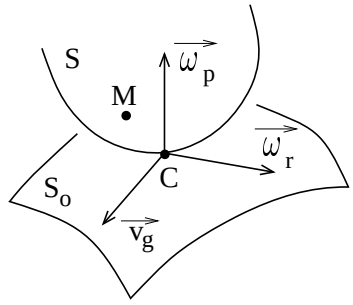


FIGURE 49

REMARQUE. Pour un solide mobile S en contact ponctuel avec un solide fixe S_0, la vitesse du point de contact C (appartenant à S) s'appelle **vitesse de glissement** $\overrightarrow{v}_g$, et les composantes $\overrightarrow{\omega}_r$ et $\overrightarrow{\omega}_p$ de $\overrightarrow{\omega}$ respectivement parallèle et orthogonale au plan de contact sont les **vitesses angulaires de roulement et de pivotement** (figure 49). Pour tout point M de S : $\overrightarrow{v}_M = \overrightarrow{v}_g + \overrightarrow{\omega} \wedge \overrightarrow{CM}$.

▷ Grandeurs mécaniques relatives à un solide

On a vu que, pour un système quelconque, la quantité de mouvement totale est reliée de façon simple à la vitesse du centre de masse : $\overrightarrow{P} = m\,\overrightarrow{v}_G$. De même pour un solide il est possible d'exprimer le moment cinétique et l'énergie cinétique en fonction de $\overrightarrow{v}_G$ et de $\overrightarrow{\omega}$. Les différentes formules s'obtiennent en particularisant un point A du solide et en écrivant pour tous les autres points M_i : $\overrightarrow{v}_i = \overrightarrow{v}_A + \overrightarrow{\omega} \wedge \overrightarrow{AM_i}$. En général on choisit pour A, soit un point fixe ($\overrightarrow{v}_A = 0$) lorsque le solide est attaché par ce point, soit le c.d.m. $G \equiv A$ (alors $\sum_i m_i \overrightarrow{AM_i} = 0$).

Dans ces deux cas le **moment cinétique** $\overrightarrow{J}_A = \sum_i \overrightarrow{AM_i} \wedge m_i\big(\overrightarrow{v}_A + \overrightarrow{\omega} \wedge \overrightarrow{AM_i}\big)$ dépend linéairement de $\overrightarrow{\omega}$ (utilisation du double produit vectoriel) :

$$\overrightarrow{J}_A = \sum_i m_i\, \overrightarrow{AM_i}^2\, \overrightarrow{\omega} - \sum_i m_i \left(\overrightarrow{AM_i} \cdot \overrightarrow{\omega}\right) \overrightarrow{AM_i} \ .$$

Il est naturel de l'écrire dans un référentiel $(AXYZ)$ lié au solide car alors les composantes X_i, Y_i et Z_i de $\overrightarrow{AM_i}$ sont indépendantes du temps. On a par exemple $J_{A,X} = \sum_i m_i\big((X_i^2+Y_i^2+Z_i^2)\,\omega_X - X_i(X_i\omega_X + Y_i\omega_Y + Z_i\omega_Z)\big)$ et plus généralement $\overrightarrow{J}_A = \mathbf{I}_A\,\overrightarrow{\omega}$ ou :

$$\begin{pmatrix} J_{A,X} \\ J_{A,Y} \\ J_{A,Z} \end{pmatrix} = \begin{pmatrix} \sum_i m_i(Y_i^2+Z_i^2) & -\sum_i m_i X_i Y_i & -\sum_i m_i X_i Z_i \\ -\sum_i m_i X_i Y_i & \sum_i m_i(X_i^2+Z_i^2) & -\sum_i m_i Y_i Z_i \\ -\sum_i m_i X_i Z_i & -\sum_i m_i Y_i Z_i & \sum_i m_i(X_i^2+Y_i^2) \end{pmatrix} \begin{pmatrix} \omega_X \\ \omega_Y \\ \omega_Z \end{pmatrix} \ .$$

La matrice d'inertie $\mathbf{I}_A$ étant symétrique, il existe des axes orthogonaux dits **axes principaux d'inertie** en A, tels qu'elle soit diagonale (*cf.* section 4.2.3).

L'**énergie cinétique** $E_C = \sum_i \dfrac{1}{2}\, m_i(\overrightarrow{v}_A + \overrightarrow{\omega} \wedge \overrightarrow{AM_i})^2$ prend elle aussi une forme simple dans ces deux cas à cause de la disparition des termes linéaires en $\overrightarrow{\omega}$. Comme $(\overrightarrow{\omega} \wedge \overrightarrow{AM_i})^2 = \omega^2\, \overrightarrow{AM_i}^2 - (\overrightarrow{\omega} \cdot \overrightarrow{AM_i})^2$ on vérifie que :

$$E_C = \frac{1}{2}\, \overrightarrow{J}_A \cdot \overrightarrow{\omega} \quad \text{(pour } A \text{ fixe)} \quad \text{ou} \quad E_C = \frac{1}{2}\, m\, \overrightarrow{v}_G^{\,2} + \frac{1}{2}\, \overrightarrow{J}_G \cdot \overrightarrow{\omega} \quad \text{(pour } A \equiv G) \ .$$

Une autre grandeur utile est la **puissance des forces** appliquées $\sum_i (\overrightarrow{v}_A + \overrightarrow{\omega} \wedge \overrightarrow{AM_i}) \cdot \overrightarrow{f}_i$ qui, quel que soit le point A du solide, s'écrit :

$$\mathcal{P} = \overrightarrow{v}_A \cdot \overrightarrow{F} + \overrightarrow{\omega} \cdot \overrightarrow{\Gamma}_A \qquad (\overrightarrow{F} = \sum_i \overrightarrow{f}_i \ \ ; \ \ \overrightarrow{\Gamma}_A = \sum_i \overrightarrow{AM_i} \wedge \overrightarrow{f}_i) \ .$$

Dans $\mathcal{P}$ seules les forces extérieures appliquées sont à considérer, car les forces intérieures $\overrightarrow{f}_{i\to j} = -\overrightarrow{f}_{j\to i}$, supposées parallèles à $\overrightarrow{M_iM_j}$, ne contribuent, ni à la force résultante $\overrightarrow{F}$, ni au moment résultant des forces $\overrightarrow{\Gamma}_A$ ($\overrightarrow{AM_i} \wedge \overrightarrow{f}_{j\to i} + \overrightarrow{AM_j} \wedge \overrightarrow{f}_{i\to j} = \overrightarrow{M_iM_j} \wedge \overrightarrow{f}_{i\to j} = 0$).

EXEMPLE 1 : **équations d'Euler**. Le mouvement libre d'un solide dans $\mathcal{R}^{cm}$ obéit à l'équation $\frac{\mathrm{d}}{\mathrm{d}t}(\overrightarrow{J}_G^{cm} = I_X\omega_X\hat{X}+I_Y\omega_Y\hat{Y}+I_Z\omega_Z\hat{Z}) = 0$, où GX, GY et GZ sont les axes principaux d'inertie en G. En écrivant $\frac{\mathrm{d}\hat{X}}{\mathrm{d}t} = \overrightarrow{\omega}\wedge\hat{X}$, *etc.* avec $\overrightarrow{\omega} = \omega_X\hat{X}+\omega_Y\hat{Y}+\omega_Z\hat{Z}$, on obtient les E.D. d'Euler :

$$I_X\dot{\omega}_X = (I_Y - I_Z)\,\omega_Y\,\omega_Z \quad , \quad I_Y\dot{\omega}_Y = (I_Z - I_X)\,\omega_Z\,\omega_X \quad , \quad I_Z\dot{\omega}_Z = (I_X - I_Y)\,\omega_X\,\omega_Y \ .$$

Les rotations à vitesse angulaire constante autour des axes GX, GY et GZ sont des solutions particulières. Si on perturbe une telle solution, par exemple $\omega_X = \omega$, $\omega_Y = \omega_Z = 0$, on obtient facilement $\ddot{\omega}_{Y,Z} = \frac{(I_Z - I_X)(I_X - I_Y)}{I_Y I_Z}\,\omega^2\,\omega_{Y,Z}$, en se limitant à l'ordre le plus bas dans les perturbations ω_Y et ω_Z. Il y a instabilité de la rotation autour de l'axe GX si $I_Y < I_X < I_Z$ (les rotations autour de GY et GZ étant alors stables).

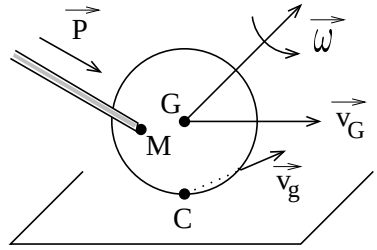


FIGURE 50

EXEMPLE 2 : **mouvement d'une boule de billard** (figure 50). Le but est d'expliquer la trajectoire,en général parabolique, du c.d.m. de la boule et le rôle des conditions initiales(façon dont la boule est frappée avec la queue). Rappelons que pour une boule de rayon r on a $\overrightarrow{J}_G = I\overrightarrow{\omega}$, avec $I = \frac{2}{5}mr^2$. La cinématique de ce problème se résume à la relation

$$\overrightarrow{v}_G = \overrightarrow{v}_g + \overrightarrow{\omega}\wedge\overrightarrow{CG} \qquad (\overrightarrow{v}_g \text{ vitesse de glissement}) \ ,$$

et la dynamique aux équations

$$m\,\overrightarrow{a}_G = m\big(\tfrac{\mathrm{d}\overrightarrow{v}_g}{\mathrm{d}t} + \tfrac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t}\wedge\overrightarrow{CG}\big) = \overrightarrow{F} \quad \text{et} \quad \tfrac{\mathrm{d}\overrightarrow{J}_G}{\mathrm{d}t} = I\,\tfrac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t} = \overrightarrow{GC}\wedge\overrightarrow{F} \ ,$$

qui conduisent à : $m\frac{\mathrm{d}\overrightarrow{v}_g}{\mathrm{d}t} = \overrightarrow{F}\big(1+\frac{mr^2}{I}\big)$; la force totale $\overrightarrow{F}$ colinéaire à $\overrightarrow{a}_G$ est parallèle à la table et donc orthogonale à $\overrightarrow{GC}$.

Au cours du mouvement et tant que la boule glisse $\overrightarrow{F}$ se réduit à une force de frottement et $\frac{\mathrm{d}\overrightarrow{v}_g}{\mathrm{d}t}$ parallèle à $\overrightarrow{F}$ l'est donc aussi à $\overrightarrow{v}_g$; $\overrightarrow{v}_g$ et $\overrightarrow{F}$ gardent donc une direction commune fixe. De plus comme pour un frottement solide $|\overrightarrow{F}|$ est constant, le vecteur $\overrightarrow{F}$ lui même est constant : le mouvement de G est alors uniformément accéléré et sa trajectoire parabolique. Bien sûr au bout d'un certain temps, $\overrightarrow{v}_g$ finit par s'annuler et la boule roule alors sans glissement tout en continuant à pivoter.

A l'origine du mouvement il faut une percussion $\overrightarrow{P}$: force $\overrightarrow{f}$ très grande et très brève (telle que $\int\overrightarrow{f}\,\mathrm{d}t = \overrightarrow{P}$) exercée en un point M à la hauteur h. Des équations $m\overrightarrow{a}_G = \overrightarrow{f}+\overrightarrow{N}$ et $I\,\frac{\mathrm{d}\overrightarrow{\omega}}{\mathrm{d}t} = \overrightarrow{GM}\wedge\overrightarrow{f}$ (la réaction normale $\overrightarrow{N}$ du billard assurant que $\overrightarrow{f}+\overrightarrow{N}$ est parallèle à la table), on déduit les conditions initiales $m\,\overrightarrow{v}_G = \overrightarrow{P}_{\parallel}$, $I\,\overrightarrow{\omega} = \overrightarrow{GM}\wedge\overrightarrow{P}$ et $\overrightarrow{v}_g = \overrightarrow{v}_G - \overrightarrow{\omega}\wedge\overrightarrow{CG}$ (la boule étant au repos avant la percussion). Si $\overrightarrow{P}$ n'est pas parallèle à la table ($\overrightarrow{P} \neq \overrightarrow{P}_{\parallel}$), on vérifie que $\overrightarrow{v}_g$ et $\overrightarrow{v}_G$ diffèrent en direction. Si $\overrightarrow{P}$ est parallèle à la table, $\overrightarrow{v}_g = \frac{\overrightarrow{P}}{m}\left(1-\frac{5}{2}\,\frac{r(h-r)}{r^2}\right)$ est du sens de $\overrightarrow{v}_G$ si $h < \frac{7}{5}r$ ("rétro") et de sens contraire si $h > \frac{7}{5}r$ ("coulé").

3.4 SYSTÈMES PHYSIQUES POSSÉDANT DES SYMÉTRIES

L'"argument de symétrie", analogue à l'"argument dimensionnel" discuté en 1.3.2, est très utilisé en physique pour simplifier la description mathématique d'un système présentant des symétries ; mais il est souvent mal justifié à cause de la confusion fréquente entre symétrie du système et symétrie (invariance) des lois le régissant (*cf.* section 3.1.1).

3.4.1 Schéma général ; principe de Curie

▷ Principe de Curie

Soient $\underline{x}$ les paramètres caractérisant un système ; par exemple pour une distribution de charges, il s'agit des positions $\overrightarrow{r}_i$ et des valeurs q_i des charges (ou de la fonction densité $\rho(\overrightarrow{r})$). Soit $\underline{y}$ une grandeur relative à ce système, par exemple le moment dipolaire $\overrightarrow{p} = \Sigma q_i \overrightarrow{r}_i$, le champ électrique $\overrightarrow{E}(\overrightarrow{r})$ créé au point $\overrightarrow{r}$, *etc.* On a vu que l'invariance des lois par rapport à une symétrie g signifie que, au système transformé dont les paramètres s'écrivent symboliquement $\underline{x}' = \mathrm{g}\underline{x}$, est associée la grandeur $\underline{y}' = \mathrm{g}\underline{y}$. Pour la distribution de charges et une rotation R, on a : $\overrightarrow{r}_i{}' = \mathrm{R}\overrightarrow{r}_i$, $q_i' = q_i$ et $\overrightarrow{p}\,' = \mathrm{R}\overrightarrow{p}$, $\overrightarrow{E}'(\mathrm{R}\overrightarrow{r}) = \mathrm{R}\overrightarrow{E}(\overrightarrow{r})$, *etc.* Si g est de plus une symétrie du système, alors par définition $\underline{x}' = \mathrm{g}\underline{x} = \underline{x}$ et donc $\underline{y}' = \mathrm{g}\underline{y} = \underline{y}$. On en déduit que les grandeurs relatives à un système ou créées ("causées") par lui possèdent au moins les symétries du système (Principe de Curie).

EXEMPLES. Si une distribution de charges ou de courants admet une symétrie de révolution d'axe Oz, son moment dipolaire électrique $\overrightarrow{p}$ ou magnétique $\overrightarrow{\mu}$ est porté par Oz ; si elle admet un plan P de symétrie, $\overrightarrow{p}$ doit être contenu dans P, tandis que $\overrightarrow{\mu}$ (pseudovecteur) doit être perpendiculaire à P, seule possibilité pour qu'ils soient inchangés par l'opération S_P (figure 51).

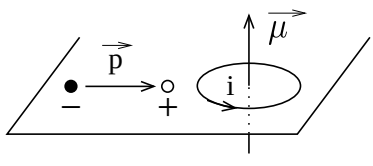


FIGURE 51

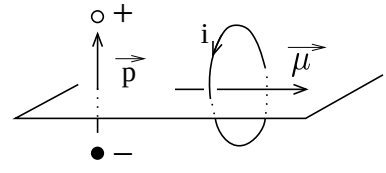


FIGURE 52

▷ Généralisation

Si la grandeur $\underline{y}$ dépend linéairement des paramètres $\underline{x}$ du système, et si l'opération g a pour effet de multiplier $\underline{x}$ par un facteur constant C (par exemple -1 pour une **antisymétrie**), alors la grandeur $\underline{y}$ est elle aussi multipliée par C par l'opération g :

$$\underline{x}' = \mathrm{g}\underline{x} = C\underline{x} \Longrightarrow \underline{y}' = \mathrm{g}\underline{y} = C\underline{y}\ .$$

EXEMPLES (figure 52). Si une distribution de charges ou de courants admet un plan P d'antisymétrie, $\overrightarrow{p}$ doit être perpendiculaire à P, tandis que $\overrightarrow{\mu}$ doit être contenu dans P, seule possibilité pour qu'ils soient multipliés par $C = -1$ dans l'opération S_P.

REMARQUE. Le groupe de symétrie de $\underline{y}$ peut être plus grand que celui de $\underline{x}$. Ainsi de nombreuses propriétés macroscopiques d'un cristal à symétrie cubique (axes x, y, z) sont celles d'un milieu isotrope. De même les moments d'inertie d'un cube homogène par rapport à des axes passant par son centre sont tous égaux comme pour une sphère. En effet considérons par exemple les quantités $I_x = \Sigma m_i(y_i^2 + z_i^2)$, $I_y = \Sigma m_i(x_i^2 + z_i^2)$ et $I_{xy} = \Sigma m_i x_i y_i$. Pour le cube tourné de $\frac{\pi}{2}$ autour de Oz, on a $x'_i = -y_i$ et $y'_i = x_i$, et donc $I'_x = I_y$, $I'_y = I_x$ et $I'_{xy} = -I_{xy}$. Comme cette rotation est une symétrie, on en déduit $I_x = I_y$ et $I_{xy} = 0$ et donc, pour tout vecteur unitaire $\hat{n}$ du plan xOy, $I_{\hat{n}} = \Sigma m_i \left((x_i \cos\theta + y_i \sin\theta)^2 + z_i^2\right) = I_x$.

3.4.2 Symétries de translation ; lois de Descartes et généralisations

L'obtention de ces lois donne un exemple non trivial d'application de l'invariance par des translations d'espace et de temps. Considérons **deux milieux linéaires, homogènes et stationnaires séparés par un plan infini** P (xOy), et une source qui émet une onde plane incidente $e^{-i(\omega_0 t - \overrightarrow{k}_0 \cdot \overrightarrow{r})}$. Soit $A\, e^{-i(\omega t - \overrightarrow{k} \cdot \overrightarrow{r})}\, e^{i\varphi}$ une onde plane réfléchie ou transmise, issue de l'onde incidente. Si l'onde incidente est retardée de τ et translatée de $\overrightarrow{a}$ parallèlement au plan P (pour laisser P inchangé), elle devient $e^{-i\left(\omega_0(t-\tau) - \overrightarrow{k}_0 \cdot (\overrightarrow{r} - \overrightarrow{a})\right)}$ et est donc multipliée par $C = e^{i(\omega_0 \tau - \overrightarrow{k}_0 \cdot \overrightarrow{a})}$. La même opération effectuée sur l'onde réfléchie ou transmise la multiplie par le facteur $e^{i(\omega\tau - \overrightarrow{k} \cdot \overrightarrow{a})}$, qui doit être aussi égal à C quels que soient τ et $\overrightarrow{a}$. On en déduit les relations de continuité :

$$\omega = \omega_0 \quad \text{et} \quad \overrightarrow{k}_{\parallel} = \overrightarrow{k}_{0\parallel} \quad \text{(loi de Descartes pour les ondes)} .$$

Donc les plans d'incidence, de réfraction et de réflexion, définis par $\overrightarrow{k}_{0\parallel}$ et la normale $\hat{n}$ au plan P, sont confondus. Les ondes ont mêmes périodicités temporelle $\dfrac{2\pi}{\omega_0} = \dfrac{2\pi}{\omega}$ et spatiale $\dfrac{2\pi}{k_{0x}} = \dfrac{2\pi}{k_x}$ en x (ou en y), et "balayent" l'axe des x à une même vitesse $\dfrac{\omega}{k_{0x}} = \dfrac{\omega}{k_x}$ (idem pour l'axe des y). Ces résultats restent valables si il y a plusieurs ondes réfléchies ou transmises, si les milieux sont anisotropes, si l'onde transmise est évanescente (k_z imaginaire pur).

Pour un **réseau plan infini à une dimension**, par exemple des traits parallèles à Oy distants de a, le plan n'est invariant que par les translations de vecteur $\overrightarrow{a} = a_y\hat{y} + na\hat{x}$ (a_y quelconque, $n \in \mathbb{Z}$). La condition précédente sur les vecteurs d'onde devient $(\overrightarrow{k} - \overrightarrow{k}_0) \cdot (a_y\hat{y} + na\hat{x}) = 0$ modulo 2π. On en déduit la formule des réseaux valable aussi bien en transmission (figure 53) qu'en réflexion :

$$\overrightarrow{k}_{\parallel} = \overrightarrow{k}_{0\parallel} + p\,\frac{2\pi}{a}\,\hat{x} \quad (p \in \mathbb{Z}) .$$

Pour un **cristal** (à l'échelle de l'Angström tout cristal macroscopique est quasi infini), l'invariance n'a lieu que pour les translations $\overrightarrow{t}_{(m,n,p)} = m\,\overrightarrow{a} + n\,\overrightarrow{b} + p\,\overrightarrow{c}$. Il faut donc $(\overrightarrow{k} - \overrightarrow{k}_0) \cdot \overrightarrow{a} = (\overrightarrow{k} - \overrightarrow{k}_0) \cdot \overrightarrow{b} = (\overrightarrow{k} - \overrightarrow{k}_0) \cdot \overrightarrow{c} = 0$ modulo 2π soit (*cf.* section 3.2.2) :

$$\begin{aligned} \overrightarrow{k} - \overrightarrow{k}_0 &= 2\pi\,\overrightarrow{\sigma}_{(h,k,l)} \quad (h,k,l \text{ quelconques}) \\ &= 2\pi\eta\,\overrightarrow{\sigma}_{(h,k,l)} \quad (h,k,l \text{ premiers entre eux},\ \eta \in \mathbb{Z}) . \end{aligned}$$

Les ondes incidente et diffractée (en dehors du cristal) ayant même longueur d'onde, on a la condition supplémentaire $|\overrightarrow{k}| = |\overrightarrow{k}_0|$ (une contrainte qui n'apparaissait pas dans les exemples précédents, k_z n'étant pas fixé). La figure 54 montre que $\overrightarrow{k}$ et $\overrightarrow{k}_0$ doivent être symétriques par rapport aux plans réticulaires, comme si il y avait réflexion sur ces plans avec un angle d'incidence θ tel que :

$$\sin\theta = \frac{2\pi\eta|\overrightarrow{\sigma}_{(h,k,l)}|}{2|\overrightarrow{k}|} = \eta\,\frac{\lambda}{2d_{(h,k,l)}} \quad (\textbf{condition de Bragg}) \ .$$

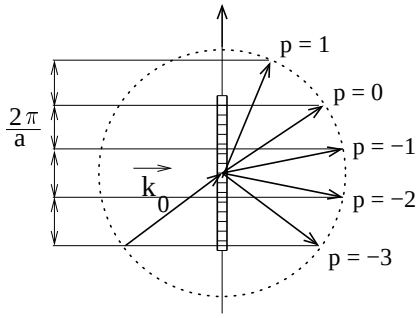


FIGURE 53

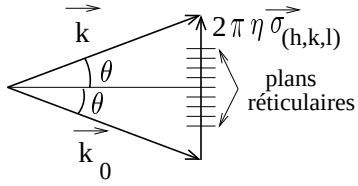


FIGURE 54

REMARQUE. Les **lois de conservation en physique quantique** de la quantité de mouvement, de l'énergie, et du moment cinétique (pour un système isolé), ont une origine semblable. En effet, d'une part la correspondance entre états quantiques à des instants différents est linéaire, et d'autre part un état ayant une valeur bien définie de $\overrightarrow{p} = \hbar\overrightarrow{k}$, ou $E = \hbar\omega$, ou $J_z = m\hbar$, est par définition multiplié respectivement par

$$C = \exp(-i\overrightarrow{k}\cdot\overrightarrow{a}) \ \text{ ou } \ C = \exp(i\omega\tau) \ \text{ ou } \ C = \exp(-im\varphi) \ ,$$

dans une translation $\overrightarrow{a}$, ou un retard τ, ou une rotation d'axe z et d'angle φ (*cf.* sections 4.4).

3.4.3 Symétries de rotation et symétries discrètes ; applications en électromagnétisme et en acoustique

Les formules reliant les symétries de la source à celles de ses effets sont applicables aux champs en tenant compte de la loi de transformation de ces derniers : $f'(\mathrm{g}\overrightarrow{r}) = \mathrm{g}f(\overrightarrow{r})$. Si l'opération g multiplie par C les paramètres de la source dont dépend le champ, on a $f' = Cf$ et donc :

$$Cf(\mathrm{g}\overrightarrow{r}) = \mathrm{g}f(\overrightarrow{r}) \quad (C = 1 \text{ pour une symétrie et } C = -1 \text{ pour une antisymétrie}) \ .$$

EXEMPLE 1. La figure 55 montre, dans les cas où g est une symétrie de rotation d'axe Oz, comment sont reliés les champs électrique $\overrightarrow{E}$ au point $M(\overrightarrow{r})$ et son transformé $M'(\mathrm{g}\overrightarrow{r})$: les composantes z, radiale et orthoradiale ne dépendent pas de θ.

EXEMPLE 2. La figure 56a montre le cas d'une symétrie plane S_P. En particulier en un point du plan P le champ $\overrightarrow{E}$ ("vrai" vecteur) est dans le plan et $\overrightarrow{B}$ (pseudovecteur) est perpendiculaire à P. Ces résultats s'échangent si le plan est un plan d'antisymétrie (figure 56b).

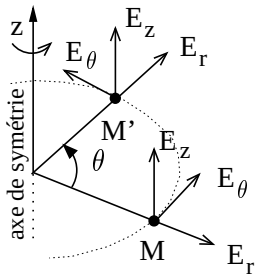


FIGURE 55

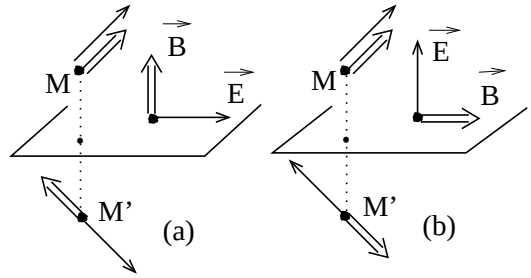


FIGURE 56

EXEMPLE 3. Si Oz est un axe de symétrie de rotation pour le système qui crée les champs et si de plus tout plan P contenant Oz est un plan de symétrie (resp. d'antisymétrie) seuls E_z, E_r et B_θ (resp. E_θ, B_z et B_r) sont non nuls. Si le plan $z = 0$ est un plan de symétrie (resp. d'antisymétrie) du système alors en deux points M et M' symétriques par rapport à ce plan les composantes E_r, E_θ et B_z (resp. E_z, B_r et B_θ) sont égales et les composantes E_z, B_r et B_θ (resp. E_r, E_θ et B_z) sont opposées. Le lecteur est encouragé à trouver des exemples physiques pour chacun de ces cas et à dessiner les figures correspondantes. C'est sur la base de ces arguments de symétrie qu'on montre "sans calcul" (en se rappelant que $\overrightarrow{B}$ est un pseudovecteur) qu'une **antenne dipolaire** électromagnétique (figure 57) ne rayonne pas dans la direction de son axe ($\overrightarrow{E} \wedge \overrightarrow{B} = 0$, car $\overrightarrow{B} = 0$, sur l'axe), ou qu'il n'y a pas de rayonnement monopolaire en électromagnétisme car pour un système à symétrie sphérique $\overrightarrow{B} = 0$ en tout point.

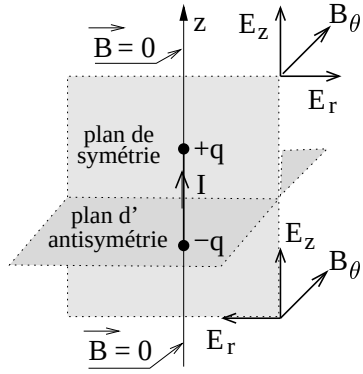


FIGURE 57

EXEMPLE 4. En **acoustique** où l'analogue du vecteur de Poynting est $P'\overrightarrow{v}$ (produit de la surpression par la vitesse du fluide), on montre pareillement qu'un dipole ne rayonne pas d'onde sonore dans une direction perpendiculaire au dipole (pas de composante du "vrai" vecteur $\overrightarrow{v}$ dans le plan d'antisymétrie du dipole), et qu'un système à symétrie sphérique (pour lequel $\overrightarrow{v}$ est radial) rayonne également dans toutes les directions (rayonnement monopolaire).

Chapitre 4

Calcul et physique linéaires ; relativité ; quantique

Le calcul linéaire trouve ses origines dans l'étude des systèmes d'équations linéaires et leur résolution par les techniques de déterminants et de transformations linéaires (à partir de 1750). Ces techniques appliquées également à la diagonalisation des formes quadratiques ont conduit vers 1850 au calcul matriciel (Cayley), ainsi qu'à la notion plus abstraite d'**espace vectoriel (E.V.)**. Son extension aux E.V. de fonctions (Hilbert 1912), le formalisme matriciel étant généralisé aux opérateurs, a joué un rôle essentiel pour la mise en forme mathématique, entre 1922 et 1927 (Schrödinger, Heisenberg, Born, *etc.*), de la "physique des quanta". Par ailleurs, les travaux initiés à partir de 1870 par Klein et Lie sur les groupes et leurs représentations linéaires (par des groupes de matrices agissant sur des E.V.), ont reçu de nombreuses applications en physique quantique à partir de 1930, l'exemple le plus simple étant l'action des rotations sur les états de spin 1/2.

Bien que d'un usage relativement tardif en physique classique, le calcul matriciel est là aussi très utile. Il permet de décrire, dans l'approximation linéaire, les relations entre grandeurs additives, que ce soient des vecteurs de l'espace ordinaire (étude des déformations, relation entre $\overrightarrow{\omega}$ et $\overrightarrow{J}$ en mécanique du solide, entre $\overrightarrow{D}$ et $\overrightarrow{E}$ en électromagnétisme...) ou des vecteurs d'E.V. plus abstraits (quadrivecteurs en relativité d'Einstein, E.V. des rayons lumineux en optique matricielle de Gauss...). Le lecteur trouvera d'autres applications inportantes aux chapitres 5 (analyse des signaux) et 6 (étude des systèmes dynamiques linéaires dans l'espace de phase).

4.1 ESPACES VECTORIELS

4.1.1 Définitions ; changements de bases ; applications linéaires

Rappelons qu'un espace vectoriel (E.V.) est un ensemble E d'éléments, les vecteurs $\underline{x}$, muni de deux lois : une loi d'addition, qui fait de E un groupe commutatif avec un élément neutre $\underline{0}$, et une loi de multiplication par des scalaires λ réels (E.V. réel) ou complexes (E.V. complexe). Ces deux lois sont distributives l'une par rapport à l'autre : $(\lambda_1 + \lambda_2)\,\underline{x} = \lambda_1 \underline{x} + \lambda_2 \underline{x}$ et $\lambda\,(\underline{x}_1 + \underline{x}_2) = \lambda \underline{x}_1 + \lambda \underline{x}_2$. Le vecteur $\underline{0} = 0 \times \underline{x}$ est simplement noté 0.

▷ Vecteurs indépendants ; bases

p vecteurs sont indépendants si l'égalité $\lambda_1 \underline{x}_1 + \cdots + \lambda_p \underline{x}_p = 0$ implique $\lambda_1 = \cdots = \lambda_p = 0$. Cette notion d'indépendance permet de définir d'une part la **dimension** de E, $\dim E$, qui est le nombre maximum n de vecteurs indépendants dans E (ou de **degrés de liberté** dans E), et d'autre part d'introduire la notion de base, qui est un ensemble discret de vecteurs indépendants $(\underline{e}_1, \cdots, \underline{e}_n)$ tels que tout $\underline{x} \in E$ s'écrit de manière unique :

$$\underline{x} = \sum_i x_i \, \underline{e}_i \quad (x_i \text{composantes ou coordonnées de } \underline{x} \text{ dans la base}) \ .$$

On dit aussi que la base $\{\underline{e}_i\}$ engendre E. $\dim E$ peut être infinie ; c'est le cas de l'E.V. des polynômes $P(x)$ engendré par la base des monomes $\{x^n, \ n \geq 0\}$, ou de celui des fonctions de période a engendré par la base des exponentielles complexes $\{e^{i2\pi n \frac{x}{a}}, \ n \in \mathbb{Z}\}$.

REMARQUES. 1) **Somme directe**. Si E_1 est un sous E.V. de E (sous ensemble de E laissé invariant par les lois ci-dessus), on appelle **sous E.V. complémentaire** E_2 tout sous E.V. tel que tout vecteur $\underline{x}$ admet une décomposition unique $\underline{x} = \underline{x}_1 + \underline{x}_2$ avec $\underline{x}_i \in E_i$. E est appelé somme directe $E = E_1 \oplus E_2$ de E_1 et E_2 et $\dim E = \dim E_1 + \dim E_2$ (**addition des degrés de liberté**). Exemple $\dim E = 2$ (base $\underline{e}_1, \underline{e}_2$) : si E_1 est engendré par $\underline{e}_1$ tout vecteur $\underline{e}_2 + \lambda \underline{e}_1$ peut engendrer E_2 (qui n'est donc pas unique ; figure 1).

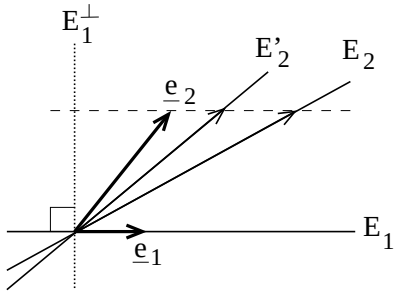


FIGURE 1

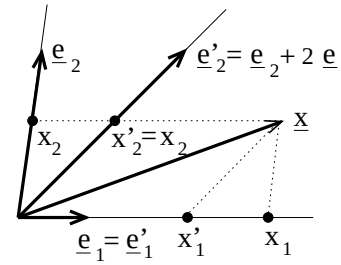


FIGURE 2

2) **Produit tensoriel**. Le produit tensoriel $E \otimes F$ de deux E.V. (E base $\{\underline{e}_i\}$ et F base $\{\underline{f}_j\}$) est l'E.V., de dimension $\dim E \times \dim F$ (**produit des nombres de degrés de liberté**), engendré par la base $\{\underline{e}_i \otimes \underline{f}_j\}$. Par exemple l'E.V. des fonctions de deux variables, périodiques de période a selon x et b selon y, est engendré par la base $\{e^{i2\pi n \frac{x}{a}} \, e^{i2\pi p \frac{y}{b}}\}$, $n, p \in \mathbb{Z}$. De même un champ $\overrightarrow{A}(\vec{r}, t) = A_x(\vec{r}, t) \, \hat{x} + A_y(\vec{r}, t) \, \hat{y} + A_z(\vec{r}, t) \, \hat{z}$ est un élément du produit tensoriel de $\mathbb{R}^3$, qui caractérise les degrés de liberté d'un vecteur, et d'un E.V. convenablement choisi de fonctions, qui caractérise ceux de la dépendance spatio-temporelle. On fera attention qu'un élément $\sum_{ij} x_{ij} \, \underline{e}_i \otimes \underline{f}_j$ de $E \otimes F$ (par exemple $2 \cos 2\pi \left(\frac{x}{a} + \frac{y}{b}\right) = e^{i2\pi \frac{x}{a}} \, e^{i2\pi \frac{y}{b}} + e^{-i2\pi \frac{x}{a}} \, e^{-i2\pi \frac{y}{b}}$) ne s'écrit pas, sauf exception, comme le produit tensoriel de deux vecteurs $\underline{x} \otimes \underline{y} = \sum_{ij} x_i y_j \, \underline{e}_i \otimes \underline{f}_j$ (ce qui est par contre le cas de $\cos 2\pi \frac{x}{a} \cos 2\pi \frac{y}{b}$).

▷ Changements de bases et de coordonnées

Ils sont reliés par :

$$\underline{e}'_i = \sum_j P_{ji} \, \underline{e}_j \Longleftrightarrow x_j = \sum_i P_{ji} \, x'_i \ .$$

En effet : $\underline{x} = \sum_i x'_i \, \underline{e}'_i = \sum_{ij} x'_i \, P_{ji} \, \underline{e}_j = \sum_j x_j \, \underline{e}_j$. On notera que si les $\underline{e}'_i$ sont donnés en fonction des $\underline{e}_j$, ce sont les x_j qui se déduisent simplement des x'_i ; l'expression des x'_i en fonction des x_j nécessite une opération d'inversion (*cf.* section 4.2.2).

EXEMPLE ($\dim E = 2$) : les relations $\underline{e}'_1 = \underline{e}_1$ et $\underline{e}'_2 = \underline{e}_2 + 2\underline{e}_1$ entraînent $x_1 = x'_1 + 2x'_2$ et $x_2 = x'_2$ (figure 2).

▷ Applications linéaires

Etant donnés deux E.V., E et F, une application linéaire A de E dans F associe à tout $\underline{x}$ de E un vecteur $\underline{y} = A\underline{x}$ de F de sorte qu'à $\lambda_1 \underline{x}_1 + \lambda_2 \underline{x}_2$ corresponde $\lambda_1 \, (A \, \underline{x}_1) + \lambda_2 \, (A \, \underline{x}_2)$. Dans des bases $\{\underline{e}_i\}$ et $\{\underline{f}_j\}$, l'application est décrite par la relation linéaire entre les composantes :

$$y_j = \sum_i A_{ji} \, x_i \quad \text{avec} \quad A_{ji} = \text{composante de } A \, \underline{e}_i \text{ selon } \underline{f}_j \ .$$

(On écrit $\underline{y} = A \sum_i x_i \, \underline{e}_i = \sum_i x_i \, (A\underline{e}_i) = \sum_i x_i \, (\sum_j A_{ji} \, \underline{f}_j)$.) Des exemples avec $E = F = \mathbb{R}^3$ sont : les rotations (*cf.* section 3.3.1) ; la relation entre $\overrightarrow{\omega}$ et $\overrightarrow{J}$ en mécanique du solide (*cf.* section 3.3.3) ; la relation entre le champ électrique $\overrightarrow{E}$ et la densité de courant $\vec{j} = Nq\vec{v}$ lorsque $m \dfrac{\vec{v}}{\tau} = q \, (\overrightarrow{E} + \vec{v} \wedge \overrightarrow{B}_0)$ (étude de l'effet Hall), qui s'écrit explicitement si $\overrightarrow{B}_0 \parallel \hat{z}$: $j_{x,y} = \left(1 + \dfrac{q^2 B_0^2 \tau^2}{m^2}\right)^{-1} \dfrac{Nq^2\tau}{m} \left(E_{x,y} \pm \dfrac{qB_0\tau}{m} \, E_{y,x}\right)$ et $j_z = \dfrac{Nq^2\tau}{m} \, E_z$; la relation entre $\overrightarrow{E}$ et $\overrightarrow{D}$ dans un diélectrique parfait anisotrope (*cf.* section 8.2.1).

AUTRES EXEMPLES : les évolutions $\underline{x}(t_0) \to \underline{x}(t) = F_{tt_0} \, \underline{x}(t_0)$ de systèmes régis par des E.D. linéaires (*cf.* section 6.1.1) ; les opérations de translation $f(t) \to f(t - \tau)$, de dérivation et plus généralement de filtrage sur les signaux (*cf.* section 5.2.1) ; les relations entre f.e.m et intensités $e_j = \sum_i R_{ji} \, I_i$ en électricité et entre flux magnétiques et intensités $\Phi_j = \sum_i M_{ji} \, I_i$ en magnétostatique.

Noyau et image. Pour A fixé, les $\underline{x}$ tels que $A\underline{x} = 0$ forment un sous E.V. de E, le noyau de A noté Ker$A$. L'ensemble des $A\underline{x}$ est un sous E.V. de F, l'image de A notée ImA ; sa dimension définit le **rang** de A et vérifie : $\text{rang}A = \dim E - \dim(\text{Ker}A)$.

EXEMPLES. Pour l'application $\vec{r} \to A\vec{r} = \hat{n} \wedge \vec{r}$, Ker$A$ est constitué par les vecteurs parallèles à $\hat{n}$ et ImA par ceux perpendiculaires à $\hat{n}$; rang$A = 2$. Dans un filtrage passe-bas idéal ($|\nu| < \nu_0$) de signaux de période T, Ker$A$ (resp. ImA) est engendré par les $e^{i2\pi n \frac{t}{T}}$ avec $|n| > \nu_0 T$ (resp. $|n| < \nu_0 T$).

Résolution d'équations linéaires $y_j = \sum_i A_{ji} \, x_i$ ($j = 1, \cdots, p$; $i = 1, \cdots, n$; x_i inconnues). Ces équations sont équivalentes à $\underline{y} = A\underline{x}$ (avec $\dim E = n$, $\dim F = p$). Donc il n'y a de solutions que si $\underline{y} \in \text{Im}A$. La solution est unique si $q = \dim(\text{Ker}A) = 0$, et elle dépend de q paramètres si $q \neq 0$. ($\underline{x}^{(1)}$ et $\underline{x}^{(2)}$ étant deux solutions, la relation $\underline{y} = A\underline{x}^{(1)} = A\underline{x}^{(2)}$ entraîne $\underline{x}^{(1)} - \underline{x}^{(2)} \in \text{Ker}A$.) En particulier pour les équations sans second membre $\sum_i A_{ji} \, x_i = 0$, une solution $\underline{x} \neq 0$ n'existe que si le noyau est non trivial ($q > 0$).

4.1.2 Structures métriques ; fonctions orthonormées

▷ Produit scalaire

Souvent, comme dans l'espace ordinaire, il existe dans E une structure métrique déduite d'un produit scalaire (appelé **produit hermitien** si E est complexe). Ce produit vérifie $(\underline{x}\,,\,\lambda_1\underline{y}_1+\lambda_2\underline{y}_2)=\lambda_1\,(\underline{x}\,,\,\underline{y}_1)+\lambda_2\,(\underline{x}\,,\,\underline{y}_2)$ et $(\underline{x}\,,\,\underline{y})=\overline{(\underline{y}\,,\,\underline{x})}$, ce qui implique, attention, $(\lambda\underline{x}\,,\,\underline{y})=\overline{\lambda}\,(\underline{x}\,,\,\underline{y})$. Dans le cas non dégénéré (considéré dans la suite), $\underline{x}\neq 0$ entraîne $(\underline{x}\,,\,\underline{x})=\|\,\underline{x}\,\|^2>0$ et $\hat{\underline{x}}=\underline{x}\,\|\,\underline{x}\,\|^{-\frac{1}{2}}$ est un vecteur normé ($\|\,\hat{\underline{x}}\,\|=1$). Le vecteur

$$\mathrm{P}_{\underline{x}}\,\underline{y}=(\hat{\underline{x}}\,,\,\underline{y})\,\hat{\underline{x}}$$

est la **projection orthogonale** de $\underline{y}$ sur (la direction de) $\underline{x}$; en effet, son complémentaire $\underline{y}'=\underline{y}-\mathrm{P}_{\underline{x}}\,\underline{y}$ est orthogonal à $\underline{x}$: $(\hat{\underline{x}}\,,\,\underline{y}')=0$. L'opérateur $\mathrm{P}_{\underline{x}}$ est un **projecteur**.

REMARQUE. Plus généralement, à tout sous E.V. E_1 est associé un unique sous E.V. complémentaire orthogonal $E_2=E_1^{\perp}$ tel que $\underline{x}=\underline{x}_1+\underline{x}_2$ avec $(\underline{x}_1\,,\,\underline{x}_2)=0$ et $\underline{x}_i\in E_i$ (*cf.* figure 1).

Inégalité de Schwarz. En écrivant que pour tout λ on a $(\underline{x}+\lambda\underline{y}\,,\,\underline{x}+\lambda\underline{y})=|\lambda|^2\;\|\,\underline{y}\,\|^2+2\Re\mathrm{e}\,\big(\lambda\,(\underline{x}\,,\,\underline{y})\big)+\|\,\underline{x}\,\|^2\geq 0$, et en choisissant λ de sorte que $\Re\mathrm{e}\,\big(\lambda\,(\underline{x}\,,\,\underline{y})\big)=|\lambda|\,|(\underline{x}\,,\,\underline{y})|$, on obtient :

$$|(\underline{x}\,,\,\underline{y})|\leq\|\,\underline{x}\,\|\;\|\,\underline{y}\,\|\qquad\text{(égalité si }\underline{y}=\mathrm{C}^{\mathrm{ste}}\,\underline{x})\;.$$

Application aux signaux. Dans l'E.V. des signaux complexes de carré sommable le produit scalaire est : $(f_1\,,\,f_2)=\int\overline{f_1}(t)\,f_2(t)\,\mathrm{d}t$. Pour un signal normalisé ($\|\,f\,\|=1$), on définit les dispersions en temps Δt et en pulsation $\Delta\omega$ par $(\Delta t)^2=\int(t-t_0)^2\,|f(t)|^2\,\mathrm{d}t$, où $t_0=\int t\,|f(t)|^2\,\mathrm{d}t$ est le temps moyen, et $(\Delta\omega)^2=\int\big|\big(\frac{1}{i}\,\frac{\mathrm{d}}{\mathrm{d}t}-\omega_0\big)\,f(t)\big|^2\,\mathrm{d}t$, où $\omega_0=\int\overline{f(t)}\,\frac{1}{i}\,\frac{\mathrm{df(t)}}{\mathrm{dt}}\,\mathrm{d}t$ est la pulsation moyenne (penser pour T grand à $f(t)$ proche de $T^{-\frac{1}{2}}\,e^{i\omega_0 t}$ sur $\big[-\frac{T}{2},\frac{T}{2}\big]$ et nulle ailleurs). En posant $\underline{x}=(t-t_0)\,f(t)$ et $\underline{y}=\big(\frac{1}{i}\,\frac{\mathrm{d}}{\mathrm{d}t}-\omega_0\big)\,f(t)$, et en notant que $|(\underline{x}\,,\,\underline{y})|\geq|\Im m(\underline{x}\,,\,\underline{y})|=\frac{1}{2}$ (calcul simple d'intégration par parties), on obtient :

$$\Delta t\,\Delta\omega\geq\tfrac{1}{2}\;.$$

Cette relation est l'analogue pour les signaux de l'**inégalité de Heisenberg**. Le lecteur montrera que l'égalité entraîne $f(t)\propto e^{i\omega_0 t}e^{-\frac{b}{2}\,(t-t_0)^2}$ avec b réel (*cf.* aussi section 4.4.3).

Bases orthonormées (orthogonalisation de Schmidt). Soit $\hat{\underline{e}}_1$ un vecteur normé. En choisissant un vecteur $\underline{e}$ indépendant de $\hat{\underline{e}}_1$ et un nombre λ_1, on construit une combinaison linéaire $\underline{e}_2=\underline{e}+\lambda_1\,\hat{\underline{e}}_1$ orthogonale à $\hat{\underline{e}}_1$; partant ensuite d'un vecteur $\underline{e}'$ indépendant de $\hat{\underline{e}}_1$ et $\hat{\underline{e}}_2$, on construit de même $\underline{e}_3=\underline{e}'+\lambda_1\,\hat{\underline{e}}_1+\lambda_2\,\hat{\underline{e}}_2$ orthogonal aux deux précédents. En itérant cette procédure et en normant les vecteurs on obtient une base telle que $(\hat{\underline{e}}_i\,,\,\hat{\underline{e}}_j)=\delta_{ij}$ ($\delta_{ij}=0$ si $i\neq j$ et $\delta_{ii}=1$ **symbole de Kronecker**). Dans une telle base :

$$(\underline{x}\,,\,\underline{y})=\sum_i\overline{x_i}y_i\;\big(=\sum_i x_iy_i\text{ dans le cas réel}\big)\;.$$

Si $\dim E$ est infini il faut "compléter" E par des "suites de Cauchy" de vecteurs (comme pour $\mathbb{R}$) : E est alors un **espace de Hilbert**.

▷ E.V. de fonctions

Un exemple est l'E.V. engendré par les fonctions $e^{im\varphi}$ ($m\in\mathbb{Z}$), qui forment une base orthonormée, pour le produit scalaire $(f\,,\,g)=(2\pi)^{-1}\int_0^{2\pi}\overline{f(\varphi)}\,g(\varphi)\,\mathrm{d}\varphi$, de l'E.V. des fonctions de période 2π de norme finie (*cf.* aussi séries de Fourier section 5.3.1).

AUTRES EXEMPLES. 1) On peut obtenir une base orthonormée de l'E.V. des fonctions définies sur $[-1,1]$, muni du produit scalaire $(f\,,\,g)=\int_{-1}^{1}\overline{f(x)}\,g(x)\,\mathrm{d}x$, en appliquant l'orthogonalisation de Schmidt aux fonctions indépendantes $1, x, \cdots, x^n$,... La base des polynômes ainsi obtenus coïncide, aux facteurs de normalisation près, avec les **polynômes de Legendre** $P_0(x)=1$, $P_1(x)=x$, $P_2(x)=\frac{1}{2}\,(3x^2-1)$..., qui sont définis par $(1-2tx+t^2)^{-\frac{1}{2}}=\sum_{l=0}^{\infty}t^l\,P_l(x)$, ou, de façon équivalente, par

$$\tfrac{1}{|\vec{r}-\vec{r}_0|}=\sum_{l=0}^{\infty}\tfrac{r_0^l}{r^{l+1}}\,P_l(\cos\theta)\ ,$$

avec $r_0<r$ et $\cos\theta=\hat{r}\cdot\hat{r}_0$ (**développement multipolaire** utilisé en électrostatique). D'autres polynômes, de Tchebychev (*cf.* section 11.3.1), de Hermite (*cf.* section 4.4.4) *etc.*, formant une base orthogonale pour d'autres produits scalaires, sont aussi utilisés en physique (rayonnement d'antennes, faisceaux optiques gaussiens, *etc.*) (*cf.* ouvrages spécialisés).

2) **Fonctions orthogonales sur la sphère**. Puisque $|\mathrm{d}(\cos\theta)|=\sin\theta\,\mathrm{d}\theta$, on pourrait prendre les fonctions $P_l(\cos\theta)\,e^{im\varphi}$ (avec $l\in\mathbb{N}$ et $m\in\mathbb{Z}$) comme base orthogonale, pour le produit scalaire $(f\,,\,g)=\iint\overline{f(\hat{n})}\,g(\hat{n})\,\mathrm{d}\Omega$ (avec $\hat{n}\equiv(\theta,\varphi)$ et $\mathrm{d}\Omega=\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$), de l'E.V. des fonctions $f(\theta,\varphi)$ de norme finie. Mais, comme le produit scalaire est invariant par rotation ($(f\,,\,g)=\iint\overline{f(R^{-1}\hat{n})}\,g(R^{-1}\hat{n})\,\mathrm{d}\Omega$), on préfère regrouper les fonctions de base par sous espaces laissés chacun invariant par rotation (*cf.* section 4.2.5). On démontre, et on admettra, que les fonctions $P_l^{|m|}(\cos\theta)\,e^{im\varphi}$ (avec $m=-l,-l+1,\cdots,l$ et $P_l^{|m|}(u)=(1-u^2)^{\frac{|m|}{2}}\,\frac{\mathrm{d}^{|m|}}{\mathrm{d}u^{|m|}}\,P_l(u)$), sont orthogonales et forment pour chaque l un sous E.V. invariant (auquel appartient $P_l(\cos\theta)$, cas $m=0$). Ces fonctions convenablement normalisées forment la base des **harmoniques sphériques** $Y_l^m(\theta,\varphi)$; les premières ($l=0,1,2$) s'écrivent :

$$Y_0^0=\tfrac{1}{\sqrt{4\pi}}\ \ ,\ \ Y_1^0=\sqrt{\tfrac{3}{4\pi}}\,\cos\theta\ \ ,\ \ Y_1^{\pm1}=\mp\sqrt{\tfrac{3}{8\pi}}\,\sin\theta\,e^{\pm i\varphi}\ ,$$
$$Y_2^0=\sqrt{\tfrac{5}{16\pi}}\,(3\,\cos^2\theta-1)\ \ ,\ \ Y_2^{\pm1}=\mp\sqrt{\tfrac{15}{8\pi}}\,\sin\theta\,\cos\theta\,e^{\pm i\varphi}\ \ ,\ \ Y_2^{\pm2}=\sqrt{\tfrac{15}{32\pi}}\,\sin^2\theta\,e^{\pm i2\varphi}\ .$$

Remarques. 1) A un facteur près ces fonctions sont 1, z et $x\pm iy$, $3z^2-1$, $z(x\pm iy)$ et $(x\pm iy)^2$; plus généralement $Y_l^{\pm l}\propto\sin^l\theta\,e^{\pm il\varphi}\propto(x\pm iy)^l$. 2) Les $P_l^{|m|}(u)$ ont $l-|m|$ zéros dans $]-1,1[$, et donc les fonctions $\Re\mathrm{e}Y_l^m$ et $\Im m Y_l^m$ s'annulent sur $l-|m|$ parallèles et $|m|$ cercles méridiens formant un "quadrillage" de la sphère. (*cf.* ouvrages de mécanique quantique.)

4.1.3 Formes quadratiques et antisymétriques ; volume

On suppose $\dim E=n$. Par définition une **forme p-linéaire** associe à p vecteurs $\underline{x}^{(1)},\ldots,\underline{x}^{(p)}$ un nombre $f(\underline{x}^{(1)},\ldots,\underline{x}^{(p)})$ qui dépend linéairement de chaque vecteur. Dans une base :

$$f(\underline{x}^{(1)},\ldots,\underline{x}^{(p)})=\sum_{i_1,\cdots,i_p}a_{i_1\cdots i_p}\,x_{i_1}^{(1)}\cdots x_{i_p}^{(p)}\qquad\text{avec}\ \ a_{i_1\cdots i_p}=f(\underline{e}_{i_1},\cdots,\underline{e}_{i_p})\ .$$

EXEMPLE. Le plus simple est l'E.V. des 1-formes $f(\underline{x})=\sum_i a_i\,x_i$, appelé **E.V. dual** E^* de E ; il est engendré par la base duale $\{f_i\}$ définie par $f_i(\underline{x})=x_i$. Si E est muni d'un produit scalaire et si $\{\underline{\hat{e}}_i\}$ est une base orthonormée, la relation $f_i(\underline{x})=(\underline{\hat{e}}_i\,,\,\underline{x})$ permet d'identifier E et E^*.

▷ Formes quadratiques

Ce sont des formes bilinéaires ($p=2$) symétriques réelles, appelées **métriques** lorsque les x_i sont des coordonnées d'espace ou d'espace-temps :

$$f(\underline{x},\underline{y})=\sum_{ij}a_{ij}\,x_i\,y_j\qquad(a_{ij}=a_{ji}\ ;\ x_i\text{ et }y_j\text{ réels})\ .$$

La forme est **définie positive** si $f(\underline{x},\underline{x})>0$ pour tout $\underline{x}\neq0$.

EXEMPLE PHYSIQUE (figure 3). L'énergie du circuit

$$e = \frac{1}{2}\left(\frac{q_1^2}{C_1} + \frac{q_2^2}{C_2}\right) + \frac{1}{2}\left(L_1\,\dot{q}_1^2 + L_2\,\dot{q}_2^2 + L(\dot{q}_1+\dot{q}_2)^2\right) = \frac{1}{2}\sum_{ij} q_i\,K_{ij}\,q_j + \frac{1}{2}\sum_{ij}\dot{q}_i\,M_{ij}\,\dot{q}_j$$

et la puissance dissipée $P = -\left(R_1\dot{q}_1^2 + R_2\dot{q}_2^2 + R(\dot{q}_1+\dot{q}_2)^2\right) = -\sum_{ij}\dot{q}_i\,F_{ij}\,\dot{q}_j$ sont des formes quadratiques respectivement positive et négative. Ici $E = \mathbb{R}^4$, $\underline{x} \equiv (q_1, q_2, \dot{q}_1, \dot{q}_2)$ et $M_{11} = L_1$, $M_{22} = L_2$, $M_{12} = M_{21} = L$, $K_{11} = \dfrac{1}{C_1}$, $K_{22} = \dfrac{1}{C_2}$, $K_{12} = K_{21} = 0$, $F_{11} = R_1 + R$, $F_{22} = R_2 + R$, $F_{12} = F_{21} = R$. Le bilan

$$\frac{\mathrm{d}e}{\mathrm{d}t} = \sum_{ij}(\dot{q}_i\,K_{ij}\,q_j + \dot{q}_i\,M_{ij}\,\ddot{q}_j) = P$$

conduit à $\sum_{ij}\dot{q}_i\left(M_{ij}\,\ddot{q}_j + F_{ij}\,\dot{q}_j + K_{ij}\,q_j\right) = 0$, ou, puisque les $\dot{q}_i$ peuvent être pris à tout instant quelconques, aux équations d'**oscillateurs couplés** :

$$\sum_j M_{ij}\,\ddot{q}_j = -\sum_j K_{ij}\,q_j - \sum_j F_{ij}\,\dot{q}_j\ .$$

(Le lecteur vérifiera que ces équations, qui généralisent celle d'un oscillateur amorti $m\ddot{x} = -kx - f\dot{x}$, sont bien celles du circuit.)

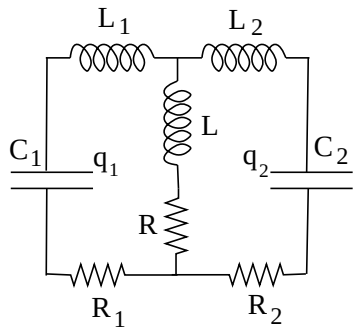


FIGURE 3

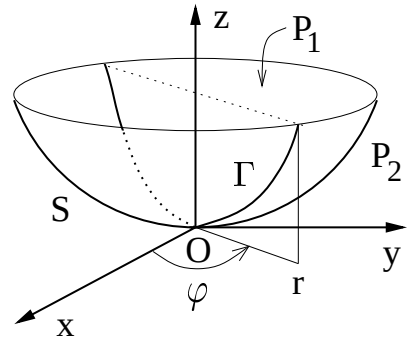


FIGURE 4

EXEMPLE GÉOMÉTRIQUE. L'équation d'une surface S tangente en O au plan (xOy) s'écrit près de O :

$$z = \frac{1}{2}\left(ax^2 + 2bxy + cy^2\right)\ ;$$

la forme est définie positive si la surface est convexe en O (figure 4). Soit $z = \dfrac{r^2}{2R(\varphi)}$ l'équation de la courbe Γ, intersection de S avec le plan ($x = r\cos\varphi$, $y = r\sin\varphi$). $R(\varphi) = \left(\dfrac{a+c}{2} + \dfrac{a-c}{2}\cos 2\varphi + b\sin 2\varphi\right)^{-1}$ est le rayon de courbure de Γ en O (*cf.* section 3.2.1). Il est compris entre les valeurs (algébriques) extrêmes, appelées aussi **rayons de courbure principaux** :

$$R_1 = \left(\frac{a+c}{2} - \sqrt{b^2 + \left(\frac{a-c}{2} \right)^2} \right)^{-1} \quad \text{et} \quad R_2 = \left(\frac{a+c}{2} + \sqrt{b^2 + \left(\frac{a-c}{2} \right)^2} \right)^{-1} .$$

En optique, la différence $R_1^{-1} - R_2^{-1}$ relative à une surface d'onde est responsable de l'aberration d'astigmatisme (*cf.* figure 34 section 7.3.2).
On appelle usuellement **courbure de la surface** S la quantité :

$$\mathcal{C} = a + c = R^{-1}(\varphi) + R^{-1}\left(\varphi + \frac{\pi}{2}\right) = R_1^{-1} + R_2^{-1} .$$

Elle vaut $\dfrac{2}{R}$ pour une sphère, $\dfrac{1}{R}$ pour un cylindre et peut être nulle si $R_1 = -R_2$ (surface en forme de "selle de cheval"). Elle intervient en hydrodynamique dans la **loi de Laplace** de la tension superficielle $P_1 - P_2 = T\,\mathcal{C}$ relative à la différence de pression de part et d'autre d'une surface séparant deux fluides (*cf.* cours de physique ; $P_1 > P_2$ sur la figure 4). Elle est différente de la **courbure de Gauss** introduite à la section 3.2.6, et qui vaut ici :

$$\mathcal{C}_{\scriptscriptstyle G} = ac - b^2 = R_1^{-1}\, R_2^{-1} .$$

▷ Formes p-alternées (ou antisymétriques)

Ce sont des formes p-linéaires ($p \leq n$) dont le signe change quand on échange deux vecteurs, $f(\cdots \underline{x} \cdots \underline{y} \cdots) = -f(\cdots \underline{y} \cdots \underline{x} \cdots)$, ou (c'est équivalent) qui s'annulent dès que deux vecteurs sont proportionnels. Des exemples pour $p = 2$ et $p = 3$ sont ($i < j < k$)

$$\sigma_{i,j}\,(\underline{x}, \underline{y}) = \begin{vmatrix} x_i & y_i \\ x_j & y_j \end{vmatrix} = x_i y_j - x_j y_i$$

$$\text{et } \sigma_{i,j,k}\,(\underline{x}, \underline{y}, \underline{z}) = \begin{vmatrix} x_i & y_i & z_i \\ x_j & y_j & z_j \\ x_k & y_k & z_k \end{vmatrix} = x_i(y_j z_k - y_k z_j) + y_i(z_j x_k - z_k x_j) + z_i(x_j y_k - x_k y_j) .$$

Les quantités $|\cdots|$ sont les déterminants des matrices $(\cdots)$ (*cf.* section 4.2.1). $\sigma_{i,j}\,(\underline{x}, \underline{y})$ représente l'**aire** (algébrique) du parallélogramme construit avec les projections de $\underline{x}$ et $\underline{y}$ sur le plan $(O, \underline{e}_i, \underline{e}_j)$, l'unité d'aire étant $\sigma_{i,j}\,(\underline{e}_i, \underline{e}_j) = 1$. Si E est l'espace usuel, $\sigma_{1,2,3}\,(\underline{x}, \underline{y}, \underline{z})$ est le produit mixte des vecteurs. De façon générale il y a C_n^p p-formes alternées indépendantes.

▷ Forme volume

Toute n-forme alternée est multiple de la forme définie par :

$$\text{vol}\,(\underline{x}^{(1)}, \ldots, \underline{x}^{(n)}) = \begin{vmatrix} x_1^{(1)} & \cdots & x_1^{(n)} \\ \vdots & \ddots & \vdots \\ x_n^{(1)} & \cdots & x_n^{(n)} \end{vmatrix} \stackrel{\text{def}}{=} \sum_{i_1, \ldots, i_n} \epsilon_{i_1 \cdots i_n}\, x_{i_1}^{(1)} \cdots x_{i_n}^{(n)} .$$

$\epsilon_{i_1 \cdots i_n}$ est un symbole qui est nul dès que deux indices sont égaux, et qui vaut ± 1 selon que la suite ordonnée de ses indices se déduit de la suite $\{1, \ldots, n\}$ par une permutation paire ou impaire (*cf.* remarque ci-dessous) :

$\epsilon_{P(1)\cdots P(n)} = (-1)^{|P|}$ ($|P| = 0$ si la permutation P est paire et $|P| = 1$ si elle est impaire) .

vol $(\cdots)$ est le volume de l'hyperparallélépipède construit avec les n vecteurs $\underline{x}^{(1)}, \ldots, \underline{x}^{(n)}$ et vol $(\underline{e}_1, \ldots, \underline{e}_n) = 1$. Son caractère alterné entraîne que vol $(\cdots) = 0$ si un des vecteurs s'exprime linéairement en fonction des autres et donc que vol $(\cdots) \neq 0$ est un critère d'indépendance des n vecteurs. Dans le cas de l'espace usuel ($n = 3$) le symbole ϵ_{ijk} est aussi noté e_{ijk} ; en utilisant la **convention de sommation d'Einstein**, qui consiste à sommer sur un indice apparaissant deux fois, et la correspondance ($x \equiv 1$, $y \equiv 2$, $z \equiv 3$), le produit vectoriel et le produit mixte s'écrivent :

$$(\overrightarrow{A} \wedge \overrightarrow{B})_i = \epsilon_{ijk}\, A_j\, B_k \quad ; \quad \overrightarrow{A} \cdot (\overrightarrow{B} \wedge \overrightarrow{C}) = \epsilon_{ijk}\, A_i\, B_j\, C_k \ .$$

REMARQUE. On rappelle que toute **permutation** est la composition de transpositions (échange de deux éléments), et que la parité du nombre de transpositions nécessaires ne dépend pas de la manière dont on procède ; par exemple la permutation qui fait passer de $\{1, 2, 3\}$ à $\{3, 2, 1\}$ est impaire car on peut la réaliser à l'aide d'une transposition (échange de 1 et 3) ou de trois (échange de 1 et 2 puis de 1 et 3 puis de 2 et 3), mais pas de deux.

4.2 CALCUL MATRICIEL

4.2.1 Bases du calcul matriciel ; lien avec le calcul vectoriel

▷ Définitions et règles de calcul

Rappelons qu'une matrice $p \times n$ est un tableau $\mathbf{M}$ de nombres (éléments du corps des réels, des complexes, *etc.*) à p lignes et n colonnes ; exemples (la notation étant justifiée ci-dessous) :

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \ , \ \mathbf{A} = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 4 & 0 \end{pmatrix} \ , \ \mathbf{A}^t = \begin{pmatrix} 2 & 1 \\ 0 & 4 \\ -1 & 0 \end{pmatrix} \ , \ \underline{x}^\dagger = \begin{pmatrix} \overline{x_1} & \overline{x_2} & \overline{x_3} \end{pmatrix} \ .$$

Le nombre situé à l'intersection de la ligne j et de la colonne i est l'élément M_{ji} de la matrice (exemple $A_{21} = 1$). $\underline{x}$ est une **matrice colonne** et $\underline{x}^\dagger$ est une **matrice ligne**.

On peut ajouter des matrices $p \times n$ entre elles ou les multiplier par λ : $(A+B)_{ji} = A_{ji} + B_{ji}$ et $(\lambda\, A)_{ji} = \lambda\, A_{ji}$ (structure d'E.V.). Le produit de deux matrices est défini par :

$$\mathbf{C} = \mathbf{B}\,\mathbf{A} \Longleftrightarrow C_{ki} = \sum_j B_{kj}\, A_{ji} \ .$$

Ceci implique que le nombre de colonnes de $\mathbf{B}$ soit égal au nombre de lignes de $\mathbf{A}$.

EXEMPLES : $\mathbf{A}\,\underline{x} = \begin{pmatrix} 2x_1 - x_3 \\ x_1 + 4x_2 \end{pmatrix}$, $\mathbf{A}\,\mathbf{A}^t = \begin{pmatrix} 5 & 2 \\ 2 & 17 \end{pmatrix}$, $\mathbf{A}^t\,\mathbf{A} = \begin{pmatrix} 5 & 4 & -2 \\ 4 & 16 & 0 \\ -2 & 0 & 1 \end{pmatrix}$.

La **matrice transposée** $\mathbf{A}^t$ déduite de $\mathbf{A}$ par échange des lignes et des colonnes ($(A^t)_{ij} = A_{ji}$), et la **matrice adjointe** $\mathbf{A}^\dagger$ définie par $(A^\dagger)_{ij} = \overline{A_{ji}}$ vérifient :

$$(\mathbf{A}^t)^t = (\mathbf{A}^\dagger)^\dagger = \mathbf{A} \quad , \quad \mathbf{C} = \mathbf{B}\,\mathbf{A} \Longleftrightarrow \mathbf{C}^t = \mathbf{A}^t\,\mathbf{B}^t \Longleftrightarrow \mathbf{C}^\dagger = \mathbf{A}^\dagger\,\mathbf{B}^\dagger \ ;$$

par exemple $(\mathbf{A}\,\underline{x})^t = \underline{x}^t\,\mathbf{A}^t = \begin{pmatrix} 2x_1 - x_3 & x_1 + 4x_2 \end{pmatrix}$. Quand elle est possible, la multiplication est associative ($\mathbf{C}\,(\mathbf{B}\,\mathbf{A}) = (\mathbf{C}\,\mathbf{B})\,\mathbf{A} = \mathbf{C}\,\mathbf{B}\,\mathbf{A}$) et distributive ($\mathbf{C}\,(\mathbf{A} + \mathbf{B}) = \mathbf{C}\mathbf{A} + \mathbf{C}\mathbf{B}$) ; donc si d$\mathbf{A}$ est une variation infinitésimale de $\mathbf{A}$ (c.a.d. $(\mathrm{d}A)_{ji} = \mathrm{d}(A_{ji})$), on a $\mathrm{d}(\mathbf{A}\mathbf{B}) = (\mathrm{d}\mathbf{A})\,\mathbf{B} + \mathbf{A}\,\mathrm{d}\mathbf{B}$. Par contre elle n'est pas commutative (*cf.* ci-dessus $\mathbf{A}^t\mathbf{A} \neq \mathbf{A}\mathbf{A}^t$). Si $p = n$ on appelle **commutateur** de $\mathbf{A}$ et $\mathbf{B}$ la matrice $[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} = -[\mathbf{B}, \mathbf{A}]$; elle vérifie $[\mathbf{A}, \mathbf{B}\mathbf{C}] = [\mathbf{A}, \mathbf{B}]\,\mathbf{C} + \mathbf{B}\,[\mathbf{A}, \mathbf{C}]$ et $[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] + [\mathbf{B}, [\mathbf{C}, \mathbf{A}]] + [\mathbf{C}, [\mathbf{A}, \mathbf{B}]] = 0$ (identité de Jacobi).

▷ Ecritures matricielle et vectorielle

Dans une base déterminée $\{\underline{e}_i\}$ de E, on représente tout **vecteur** $\underline{x} = \sum_i x_i\,\underline{e}_i$ par une **matrice colonne**, notée aussi $\underline{x}$, d'éléments x_i (*cf.* exemple ci-dessus) ; en particulier la matrice $\underline{e}_i$ est la colonne dont tous les éléments sont nuls, sauf celui de la ligne i qui est égal à 1. On en déduit l'écriture matricielle d'une application linéaire de E (base $\{\underline{e}_i\}$) dans F (base $\{\underline{f}_j\}$)

$$y_j = \sum_i A_{ji} x_i \Longleftrightarrow \underline{y} = \mathbf{A}\underline{x}$$

(en identifiant l'application et sa matrice représentative dans les bases choisies). La $i^{\text{ème}}$ colonne de $\mathbf{A}$ n'est autre que le produit matriciel $\mathbf{A}\underline{e}_i$ et $A_{ji} = \underline{f}_j^t \mathbf{A} \underline{e}_i$. Si $E \equiv F$, un produit scalaire de vecteurs dans une base orthonormée s'écrit :

$$(\underline{x}, \underline{y}) = \sum_i \overline{x}_i y_i = \underline{x}^\dagger \underline{y} \quad (= \underline{x}^t \underline{y} \text{ dans le cas réel}) \ ; \ (A\underline{x}, B\underline{y}) = \underline{x}^\dagger \mathbf{A}^\dagger \mathbf{B} \underline{y} \ .$$

4.2.2 Matrices $n \times n$; exponentielle ; trace ; déterminant ; inverse

▷ Matrices remarquables

1) La **matrice identité** $\mathbf{I}$ d'éléments $I_{ji} = \delta_{ji}$. Elle vérifie $\mathbf{I}\mathbf{A} = \mathbf{A}\mathbf{I} = \mathbf{A}$ et c'est la seule (à un facteur près) à commuter avec toutes les autres.
2) Les matrices réelles **symétriques** $\mathbf{S} = \mathbf{S}^t$, **antisymétriques** $\mathbf{A} = -\mathbf{A}^t$, et les matrices **orthogonales** $\mathbf{O}\mathbf{O}^t = \mathbf{O}^t\mathbf{O} = \mathbf{I}$ qui conservent le produit scalaire réel : $(O\underline{x}, O\underline{y}) = \underline{x}^t\mathbf{O}^t\mathbf{O}\underline{y} = (\underline{x}, \underline{y})$. Par exemple pour $n = 2$:

$$\mathbf{S} = \begin{pmatrix} a & b \\ b & d \end{pmatrix} ; \ \mathbf{A} = \begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix} ; \ \mathbf{O} = \mathbf{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ ou } \mathbf{S}_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} .$$

$\mathbf{R}_\theta$ correspond à une rotation d'angle θ dans le plan (xy), et $\mathbf{S}_{2\theta} = \mathbf{R}_\theta \mathbf{S}_0 \mathbf{R}_\theta^{-1}$ est une symétrie par rapport à un axe faisant l'angle θ avec Ox.
3) Les matrices complexes **hermitiennes** $\mathbf{H} = \mathbf{H}^\dagger$, **antihermitiennes** $i\mathbf{H}$ telles que $(i\mathbf{H})^\dagger = -i\mathbf{H}$, et les matrices **unitaires** $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$ qui conservent le produit scalaire hermitien : $(U\underline{x}, U\underline{y}) = \underline{x}^\dagger\mathbf{U}^\dagger\mathbf{U}\underline{y} = (\underline{x}, \underline{y})$.
4) Les **projecteurs**. $\mathbf{P} = \underline{m}\,\underline{m}^\dagger$ (avec $\underline{m}$ normé) projette le vecteur $\underline{x}$ sur la direction $\underline{m}$. On vérifie : $\mathbf{P} = \mathbf{P}^\dagger$, $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{P}_1\mathbf{P}_2 = \underline{m}_1(\underline{m}_1^\dagger\underline{m}_2)\underline{m}_2^\dagger = 0$ si $\underline{m}_1$ et $\underline{m}_2$ sont orthogonaux.

EXEMPLE. En **optique ondulatoire**, un polariseur rectiligne, ou circulaire gauche (resp. droit), est un appareil qui projette le champ complexe $\begin{pmatrix} E_x \\ E_y \end{pmatrix} e^{-i\omega t}$ sur les vecteurs $\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, ou $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$ (resp. $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$). Ces vecteurs normés représentent les états de polarisation des photons associés au champ (*cf.* section 4.4.1). Les projecteurs $\begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix}$ ou $\frac{1}{2} \begin{pmatrix} 1 & \mp i \\ \pm i & 1 \end{pmatrix}$ représentent l'"action" du polariseur.

▷ Exponentielles ; groupes continus

On démontre, et on admettra, que l'étude des groupes "continus" (rotations par exemple) se ramène à celle de sous groupes à un paramètre (rotations autour d'un axe par exemple) ; ces sous groupes font intervenir des exponentielles de matrices (générateurs des sous groupes). Comme pour les nombres, $e^{\mathbf{A}}$ est défini par la série convergente

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^n}{n!} + \cdots$$

et les matrices $\mathbf{M}(t) = e^{t\mathbf{A}}$ vérifient les propriétés :

$$e^{t_1\mathbf{A}}\, e^{t_2\mathbf{A}} = e^{(t_1+t_2)\,\mathbf{A}} \quad , \quad \frac{\mathrm{d}}{\mathrm{d}t}\, e^{t\mathbf{A}} = \mathbf{A}\, e^{t\mathbf{A}} \quad , \quad e^{t\mathbf{A}} = \lim_{n\to\infty} \Big(\mathbf{I} + t\,\frac{\mathbf{A}}{n}\Big)^n .$$

Ces matrices forment un groupe commutatif à un paramètre entièrement déterminé par son **générateur** $\mathbf{A} = \left.\dfrac{\mathrm{d}\mathbf{M}(t)}{\mathrm{d}t}\right|_{t=0}$.

EXEMPLE 1. Les rotations (actives) d'angle θ dans le plan (x, y), les transformations (passives) de Lorentz de rapidité φ dans le plan (x, ct) ou de Galilée de vitesse V dans le plan (x, t) s'écrivent respectivement : $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \exp\left[\theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right]$, $\begin{pmatrix} \cosh\varphi & -\sinh\varphi \\ -\sinh\varphi & \cosh\varphi \end{pmatrix} = \exp\left[-\varphi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right]$ et $\begin{pmatrix} 1 & -V \\ 0 & 1 \end{pmatrix} = \exp\left[-V \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right]$; le calcul de $e^{\mathbf{A}}$ est immédiat car $\mathbf{A}^2 = -\mathbf{I}$, $\mathbf{I}$ ou $\mathbf{0}$.

EXEMPLE 2. Les **rotations dans l'espace** d'angle θ autour d'un axe passant par O et de vecteur directeur $\hat{n}(\alpha, \beta, \gamma)$ s'écrivent :

$$\mathbf{R}_{\hat{n}}(\theta) = e^{\theta\,\mathbf{A}_{\hat{n}}} \quad \text{avec} \quad \mathbf{A}_{\hat{n}} = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} .$$

En effet, de $\mathbf{A}_{\hat{n}}\,\vec{r} = \hat{n} \wedge \vec{r}$ on déduit $\mathbf{R}_{\hat{n}}(\epsilon)\,\vec{r} = \vec{r} + \epsilon\,\hat{n} \wedge \vec{r}$, qui correspond bien à une rotation infinitésimale autour de $\hat{n}$. Les matrices $\mathbf{R}_{\hat{n}}(\theta)$ s'identifient aux matrices orthogonales 3×3 de déterminant 1.

EXEMPLE 3. Si $\mathbf{A} = -\mathbf{A}^t$ est réelle, les matrices $\mathbf{O}(t) = e^{t\mathbf{A}}$ sont orthogonales (en effet $(\exp t\mathbf{A})^t = \exp t\mathbf{A}^t = \exp -t\mathbf{A}$), et réciproquement (on écrit $\mathbf{O}(\epsilon)\,\mathbf{O}^t(\epsilon) = (\mathbf{I}+\epsilon\mathbf{A})\,(\mathbf{I}+\epsilon\mathbf{A}^t) = \mathbf{I}$ à l'ordre ϵ). Si $\mathbf{H} = \mathbf{H}^\dagger$ est hermitienne, les matrices $\mathbf{U}(t) = e^{-it\mathbf{H}}$ sont **unitaires** et réciproquement (démonstration analogue).

EXEMPLE 4. Les solutions de $\dot{\underline{x}} = \mathbf{A}\,\underline{x}$, équation à laquelle se ramène toute E.D.L.S. (*cf.* section 6.2.2), sont $\underline{x}(t) = e^{t\mathbf{A}}\,\underline{x}(0)$.

REMARQUE. **Groupe de Lie**. On fera attention que $e^{\mathbf{A}}\,e^{\mathbf{B}} \neq e^{\mathbf{A}+\mathbf{B}}$ si les matrices $\mathbf{A}$ et $\mathbf{B}$ ne commutent pas. Le résultat exact, dont la démonstration est non triviale, est $e^{\mathbf{A}}\,e^{\mathbf{B}} = e^{\mathbf{C}}$ avec $\mathbf{C} = \mathbf{A}+\mathbf{B}+\frac{1}{2}[\mathbf{A},\mathbf{B}]+\cdots$, où les « $\cdots$ » ne font intervenir que des matrices obtenues à partir de commutations avec $\mathbf{A}$ ou $\mathbf{B}$ de leur commutateur $[\mathbf{A},\mathbf{B}]$. On le vérifie à l'ordre deux en écrivant $e^{\mathbf{A}}\,e^{\mathbf{B}} = (\mathbf{I}+\mathbf{A}+\frac{\mathbf{A}^2}{2}+\cdots)\,(\mathbf{I}+\mathbf{B}+\frac{\mathbf{B}^2}{2}+\cdots) = \mathbf{I}+\mathbf{A}+\mathbf{B}+\frac{(\mathbf{A}+\mathbf{B})^2}{2}+\frac{\mathbf{AB}-\mathbf{BA}}{2}+\cdots$. Ce résultat a pour conséquence qu'une famille de matrices $\mathbf{M}(\underline{t})$ dépendant continûment de p paramètres $t_1,\cdots,t_p$, et reliées continûment à l'identité $\mathbf{M}(0) = \mathbf{I}$, forme un groupe (de Lie) si et seulement si les commutateurs des p générateurs $\mathbf{A}_i = \frac{\partial \mathbf{M}(\underline{t})}{\partial t_i}|_{\underline{t}=0}$ sont des combinaisons linéaires des $\mathbf{A}_j$. Par exemple pour les générateurs $\mathbf{A}_{\hat{n}}$ ci-dessus du groupe des rotations, un calcul élémentaire montre que : $[\mathbf{A}_{\hat{x}},\mathbf{A}_{\hat{y}}] = \mathbf{A}_{\hat{z}}$, $[\mathbf{A}_{\hat{y}},\mathbf{A}_{\hat{z}}] = \mathbf{A}_{\hat{x}}$ et $[\mathbf{A}_{\hat{z}},\mathbf{A}_{\hat{x}}] = \mathbf{A}_{\hat{y}}$.

▷ Trace et déterminant

On appelle **trace** de $\mathbf{A}$ le nombre

$$\mathrm{tr}\mathbf{A} = \sum_i A_{ii} \text{ (somme des éléments diagonaux)} .$$

On vérifie facilement : $\mathrm{tr}\mathbf{I} = n$, $\mathrm{tr}\mathbf{A}^t = \mathrm{tr}\mathbf{A}$ et $\mathrm{tr}\mathbf{AB} = \mathrm{tr}\mathbf{BA}$.
Le **déterminant** est défini par :

$$\det\mathbf{A} = \begin{vmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{vmatrix} = \sum_{i_1\cdots i_n} \epsilon_{i_1\cdots i_n} A_{i_1 1}\cdots A_{i_n n} .$$

Cette quantité, écrite explicitement pour $n = 2$ et $n = 3$ à la section 4.1.3, n'est autre que $\mathrm{vol}(\mathbf{A}\underline{e}_1,\cdots,\mathbf{A}\underline{e}_n)$. On en déduit que $\det\mathbf{A}$ est multiplié par λ si une colonne est multipliée par λ (donc $\det(\lambda\mathbf{A}) = \lambda^n \det\mathbf{A}$), et qu'il reste inchangé si on ajoute à une colonne une combinaison linéaire d'autres colonnes ; la condition $\det\mathbf{A} \neq 0$ caractérise l'indépendance des vecteurs "colonnes" $\mathbf{A}\underline{e}_i$. On montre que les déterminants vérifient :

$$\det\mathbf{A}^t = \det\mathbf{A} \quad , \quad \det\mathbf{AB} = \det\mathbf{A}\,\det\mathbf{B} \quad , \quad \det e^{\mathbf{A}} = e^{\mathrm{tr}\mathbf{A}} .$$

La dernière propriété résulte de : $\det\big(\mathbf{I}+\mathbf{A}/n\big)^n \simeq (1+\mathrm{tr}\mathbf{A}/n)^n$ pour n grand.

REMARQUE. Dans la somme qui définit $\det\mathbf{A}$, chaque terme non nul contient un et un seul élément A_{ji} pour j fixé. On peut donc regrouper les termes de cette somme pour développer le déterminant par rapport aux éléments d'une ligne j quelconque (mais fixée) et écrire $\det\mathbf{A} = \sum_i A_{ji}\widetilde{A}_{ji}$. Le coefficient $\widetilde{A}_{ji}$, appelé **cofacteur** de A_{ji}, est le produit par $(-1)^{i+j}$ du déterminant de la matrice obtenue en supprimant dans $\mathbf{A}$ la ligne j et la colonne i. On ramène ainsi le calcul d'un déterminant $n\times n$ à celui de déterminants $(n-1)\times(n-1)$ (*cf.* $\sigma_{i,j,k}(\underline{x},\underline{y},\underline{z})$ section 4.1.3). On a aussi $\sum_i A_{ji}\widetilde{A}_{ki} = \det\mathbf{A}\,\delta_{jk}$, car (réflexion laissée au lecteur) pour $k \neq j$ la somme est le déterminant de la matrice $\mathbf{A}$ dans laquelle la ligne k a été remplacée par la ligne j, et un déterminant est nul si deux lignes j et k sont égales, d'où : $\mathbf{A}\widetilde{\mathbf{A}}^t = \widetilde{\mathbf{A}}^t\mathbf{A} = (\det\mathbf{A})\,\mathbf{I}$ ($\widetilde{\mathbf{A}}$ matrice des cofacteurs).

▷ Inverse

La matrice inverse $\mathbf{A}^{-1}$ de $\mathbf{A}$ est telle que $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Son unicité est facile à démontrer ; la remarque ci-dessus établit sa condition d'existence $\det\mathbf{A} \neq 0$.

EXEMPLE. $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ si $\Delta = ad - bc \neq 0$.

Application aux changements de bases. Si $\underline{y} = \mathbf{A}\underline{x}$ est une relation matricielle entre vecteurs de E, un changement de base induit les changements de coordonnées $\underline{x} = \mathbf{P}\underline{x}'$, $\underline{y} = \mathbf{P}\underline{y}'$ (*cf.* section 4.1.1) ; il conduit à $\underline{y}' = \mathbf{A}'\underline{x}'$ avec $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, matrice représentative de la relation dans la nouvelle base.

4.2.3 Spectre d'une matrice $n \times n$; vecteurs propres ; exemples

Le **spectre** de $\mathbf{A}$ est l'ensemble des nombres λ pour lesquels $\mathbf{A} - \lambda\mathbf{I}$ n'est pas inversible, donc pour lesquels $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, ou encore tels que $\mathrm{Ker}(\mathbf{A} - \lambda\mathbf{I}) \neq 0$. Sa connaissance est essentielle pour mettre $\mathbf{A}$ sous une forme simple (diagonale ou de Jordan), et pour ramener toute fonction de $\mathbf{A}$ à un polynôme simple.

▷ Polynôme caractéristique de A

Défini par

$$P_c(z) \stackrel{\text{def}}{=} \det(z\mathbf{I} - \mathbf{A}) = \Pi_{i=1}^{n} (z - \lambda_i) \ ,$$

donc de la forme $z^n + a_{n-1}z^{n-1} + \cdots + a_0$, avec $(-1)^n a_0 = \det \mathbf{A} = \Pi_i \lambda_i$ et $-a_{n-1} = \mathrm{tr}\mathbf{A} = \sum_i \lambda_i$, il a pour zéros le spectre de $\mathbf{A}$. Comme $\det(\mathbf{P}\mathbf{P}^{-1}) = \det \mathbf{P} \det(\mathbf{P}^{-1}) = \det \mathbf{I} = 1$, les matrices $\mathbf{A}$ et $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, qui correspondent à la même application linéaire dans différentes bases, ont même polynôme caractéristique.

Théorème de Cayley-Hamilton. Le polynôme caractéristique est annulé par $\mathbf{A}$: $P_c(\mathbf{A}) = 0$; par exemple pour $n = 2$: $\mathbf{A}^2 - (a + d)\,\mathbf{A} + (ad - bc)\,\mathbf{I} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ avec $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

DÉMONSTRATION. Elle repose sur la remarque que dans la matrice des cofacteurs de $z\mathbf{I} - \mathbf{A}$, et donc aussi dans sa transposée notée $\mathbf{B}(z)$, le plus haut degré possible pour z est z^{n-1}. En écrivant $(z\mathbf{I} - \mathbf{A})\,\mathbf{B}(z) = P_c(z)\,\mathbf{I}$ avec $\mathbf{B}(z) = z^{n-1}\mathbf{B}_{n-1} + \cdots + \mathbf{B}_0$, et en identifiant les termes en z^p ($p = 0, \cdots, n$) dans les deux membres, il vient $a_0\mathbf{I} = -\mathbf{A}\mathbf{B}_0$, $a_1\mathbf{I} = -\mathbf{A}\mathbf{B}_1 + \mathbf{B}_0, \cdots$, $\mathbf{I} = \mathbf{B}_{n-1}$, d'où $a_0\mathbf{I} + a_1\mathbf{A} + \cdots + a_{n-1}\mathbf{A}^{n-1} + \mathbf{A}^n = 0$.

De ce théorème découle l'existence d'un **polynôme minimal** $P_m(z)$ de degré $p \leq n$, unique à un facteur près (identique à $P_c(z)$ si $p = n$), et tel que $P_m(\mathbf{A}) = 0$. On peut montrer, en divisant $P_c(z)$ par $P_m(z)$, que $P_m(z)$ possède, avec éventuellement des multiplicités moindres, tous les zéros de $P_c(z)$. Par exemple les matrices $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ et $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ ont pour polynôme minimal respectif $(z - \lambda_1)(z - \lambda_2) = P_c(z)$ si $\lambda_1 \neq \lambda_2$ ou $(z - \lambda)$ si $\lambda_1 = \lambda_2 = \lambda$, et $(z - \lambda)^2 = P_c(z)$ (vérification immédiate). L'intérêt de P_m est d'assurer que tout polynôme $P(\mathbf{A})$, ou toute série convergente de $\mathbf{A}$, s'exprime en fonction des seules matrices $\mathbf{I}, \mathbf{A}, \cdots, \mathbf{A}^{p-1}$; en effet, de la division $P(z) = P_m(z)\,Q(z) + R(z)$ on déduit explicitement $P(\mathbf{A}) = R(\mathbf{A})$.

▷ Vecteurs propres et valeurs propres ; diagonalisation

Soit λ un élément du spectre de $\mathbf{A}$. La propriété $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ entraîne que l'équation $\mathbf{A}\,\underline{m} = \lambda\,\underline{m}$ admet au moins une solution $\underline{m} \neq 0$. Le vecteur $\underline{m}$, défini à une constante multiplicative près, est appelé **vecteur propre** associé à la valeur propre λ. Si la matrice $\mathbf{A}$ possède une base de vecteurs propres

$$\mathbf{A}\,\underline{m}_i = \lambda_i\,\underline{m}_i \quad (i = 1, \cdots, n;\ \underline{m}_i \text{ vecteurs indépendants}) \ ,$$

elle devient une matrice diagonale (avec les λ_i comme éléments diagonaux et des zéros ailleurs) lorsqu'on l'écrit dans cette base et on retrouve :

$$\det \mathbf{A} = \Pi_i \, \lambda_i \quad \text{et} \quad \mathrm{tr}\mathbf{A} = \sum_i \lambda_i \ .$$

En effet, si $\mathbf{P}$ est la matrice qui a pour colonne i les composantes de $\underline{m}_i$ dans la base initiale, $\mathbf{P}^{-1}\mathbf{P} = \mathbf{I}$ entraîne que $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\left(\lambda_1\underline{m}_1 \quad \cdots \quad \lambda_n\underline{m}_n\right)$ a pour éléments $A'_{ij} = \lambda_i \, \delta_{ij}$. Une telle base existe toujours si il y a n valeurs propres différentes car des vecteurs propres relatifs à des valeurs propres différentes sont indépendants ; par exemple si $\underline{m}_2 = a\,\underline{m}_1 \neq 0$, l'application de $\mathbf{A}$ donne $\lambda_2\,\underline{m}_2 = a\lambda_1\,\underline{m}_1$, ce qui est contradictoire pour $\lambda_2 \neq \lambda_1$; on ne peut avoir de même $\underline{m}_3 = a\,\underline{m}_1 + b\,\underline{m}_2$, avec $\underline{m}_1$ et $\underline{m}_2$ indépendants et $\lambda_1 \neq \lambda_2 \neq \lambda_3$.

REMARQUES. 1) La relation $\mathbf{A}^n = (\mathbf{P}\mathbf{A}'\mathbf{P}^{-1})^n = \mathbf{P}\mathbf{A}'^n\mathbf{P}^{-1}$ simplifie le calcul de $\mathbf{A}^n$ puisque $(\mathbf{A}'^n)_{ij} = \lambda_i^n \, \delta_{ij}$ (et le calcul d'autres fonctions de $\mathbf{A}$ comme $\exp \mathbf{A}$). 2) Si $\mathbf{A}$ et $\mathbf{B}$ sont diagonalisables et si $[\mathbf{A}, \mathbf{B}] = 0$, on peut montrer qu'elles ont une base commune de vecteurs propres.

Matrice de Jordan. $\mathbf{A}$ n'est pas diagonalisable si, lorsque λ est un zéro d'ordre $q > 1$ de $P_c(z)$ (λ dégénérée q fois), l'équation $(\mathbf{A} - \lambda\mathbf{I})\,\underline{m} = 0$ ne fournit pas q vecteurs indépendants. On montre que par contre $(\mathbf{A} - \lambda\mathbf{I})^q \, \underline{m} = 0$ a q solutions indépendantes $\underline{m}_i$ qu'on peut choisir de sorte que $\mathbf{A}\underline{m}_i = \lambda\underline{m}_i$ ou $\mathbf{A}\underline{m}_i = \lambda\underline{m}_i + \underline{m}_{i-1}$. Par exemple pour $n = q = 2$ ou $n = q = 3$, et $\mathbf{A}$ non diagonalisable, $\mathbf{A}'$ peut être réduite à la forme $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ ou à l'une des deux formes $\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$ ou $\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$.

▷ Matrices remarquables

Les matrices hermitiennes $\mathbf{H} = \mathbf{H}^\dagger$ (ou symétriques réelles $\mathbf{S} = \mathbf{S}^t$), les matrices anti-hermitiennes $i\mathbf{H}$ (ou antisymétriques réelles $\mathbf{A} = -\mathbf{A}^t$), les matrices unitaires $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$ (ou orthogonales réelles $\mathbf{O}\mathbf{O}^t = \mathbf{O}^t\mathbf{O} = \mathbf{I}$) possèdent toutes une **base orthonormée de vecteurs propres**. Les valeurs propres de ces trois groupes de matrices sont respectivement **réelles, imaginaires pures, de module 1** ; les vecteurs propres sont réels pour les matrices symétriques réelles. Ces propriétés reposent sur la relation $(\underline{x}, A\underline{y}) = \underline{x}^\dagger\mathbf{A}\underline{y} = (\mathbf{A}^\dagger\underline{x})^\dagger\underline{y}$. En décomposant les vecteurs sur la base orthonormée des vecteurs propres $\underline{m}_i$ de $\mathbf{A}$, il vient $(\underline{x}, A\underline{y}) = (\sum_i X_i\underline{m}_i, A\sum_j Y_j\underline{m}_j) = \sum_i \overline{X_i}Y_i\lambda_i = \sum_i \lambda_i \, (\underline{x}, \underline{m}_i)\,(\underline{m}_i, \underline{y})$. La matrice $\mathbf{A}$ s'écrit donc aussi (décomposition spectrale de $\mathbf{A}$) :

$$\mathbf{A} = \sum_i \lambda_i \, \mathbf{P}_i \quad \text{avec } \mathbf{P}_i = \underline{m}_i \, \underline{m}_i^\dagger \text{ (projecteur sur } \underline{m}_i) \ .$$

JUSTIFICATION (on admet la diagonalisabilité). Considérons le cas $\mathbf{A} = \mathbf{A}^\dagger$; en prenant $\underline{x} = \underline{y} = \underline{m}$ vecteur propre on obtient $\overline{\lambda} = \lambda$, et en prenant $\underline{x} = \underline{m}_1$ et $\underline{y} = \underline{m}_2$ on obtient $\lambda_1(\underline{m}_1, \underline{m}_2) = \lambda_2(\underline{m}_1, \underline{m}_2)$, donc $(\underline{m}_1, \underline{m}_2) = 0$ si $\lambda_1 \neq \lambda_2$; (si $\lambda_1 = \lambda_2$ il existe des combinaisons de $\underline{m}_1$ et $\underline{m}_2$ orthogonales). Le cas $\mathbf{A} = -\mathbf{A}^\dagger$ se ramène au cas précédent car $i\mathbf{A}$ est hermitienne. Quant aux matrices unitaires, on déduit de $(U\underline{m}_1, U\underline{m}_2) = \overline{\lambda_1}\lambda_2 \, (\underline{m}_1, \underline{m}_2) = (\underline{m}_1, \underline{m}_2)$ que $|\lambda| = 1$ en posant $\underline{m}_1 = \underline{m}_2 = \underline{m}$ et que $(\underline{m}_1, \underline{m}_2) = 0$ si $\lambda_1 \neq \lambda_2$.

EXEMPLE 1. La conique $ax^2 + 2bxy + dy^2 = \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a & b \\ b & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 1$ s'écrit $\lambda_1 X^2 + \lambda_2 Y^2 = 1$ dans la base des vecteurs propres normés. Si le déterminant $ad - b^2 = \lambda_1\lambda_2$ est positif, c'est un **ellipse** de demi axes $\lambda_{1,2}^{-\frac{1}{2}}$ et donc d'aire $\pi\,(ad - b^2)^{-\frac{1}{2}}$.

EXEMPLE 2. Les matrices de rotation $\mathbf{R}_\theta$ (*cf.* section 4.2.2) ont pour valeurs propres et vecteurs propres normés $\lambda_\pm = e^{\mp i\theta}$ et $\underline{m}_\pm = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \pm i\end{pmatrix}$. Les polarisations gauche et droite en **optique ondulatoire**, et les états des photons associés, sont donc des vecteurs propres des rotations autour de l'axe de propagation z avec pour valeurs propres $e^{-i\theta}$ et $e^{i\theta}$ (*cf.* section 4.4.1).

▷ Modes propres d'oscillateurs couplés

En notation complexe, on appelle ainsi les solutions de l'équation (*cf.* exemple électrique section 4.1.3)

$$\mathbf{M}\ddot{\underline{x}} = -\mathbf{K}\underline{x}$$

du type $\underline{x}(t) = e^{i\omega t}\,\underline{m}$ (avec $\mathbf{M}$ et $\mathbf{K}$ symétriques définies positives). $\underline{m}$ et ω doivent vérifier $(\mathbf{K} - \omega^2\mathbf{M})\,\underline{m} = 0$ et $\det(\mathbf{K} - \omega^2\mathbf{M}) = 0$; ce problème est équivalent à la recherche des valeurs et vecteurs propres de la matrice symétrique $\mathbf{M}^{-\frac{1}{2}}\mathbf{K}\mathbf{M}^{-\frac{1}{2}}$. De $\underline{m}_1^t\mathbf{K}\underline{m}_2 = \omega_2^2\,\underline{m}_1^t\mathbf{M}\underline{m}_2 = (\underline{m}_2^t\mathbf{K}\underline{m}_1)^t = \omega_1^2\,\underline{m}_1^t\mathbf{M}\underline{m}_2$ on déduit, en faisant $\underline{m}_1 = \underline{m}_2 = \underline{m}$, que $\omega^2 > 0$ (ω réel), et que les moyennes temporelles des "énergies potentielle" $\frac{1}{2}\,\underline{x}^t\mathbf{K}\underline{x}$ et "cinétique" $\frac{1}{2}\,\dot{\underline{x}}^t\mathbf{M}\dot{\underline{x}}$ sont égales dans un mode propre ; si $\omega_1 \neq \omega_2$ on a : $\underline{m}_1^t\mathbf{M}\underline{m}_2 = 0$. Ces modes sont étudiés à la section 6.2.2.

▷ Calculs de perturbation

Matrices hermitiennes $\mathbf{A} = \mathbf{A}^\dagger$. 1) Si λ_i est une valeur propre (rélle) non dégénérée et si la perturbation $\delta\mathbf{A}$ de $\mathbf{A}$ est "petite", alors $\delta\underline{m}_i$ est "petit". De $\lambda_i = \underline{m}_i^\dagger\,\mathbf{A}\,\underline{m}_i$ on déduit :

$$\delta\lambda_i = \lambda_i((\delta\underline{m}_i^\dagger)\,\underline{m}_i + \underline{m}_i^\dagger\,(\delta\underline{m}_i)) + \underline{m}_i^\dagger\,(\delta\mathbf{A})\,\underline{m}_i = \underline{m}_i^\dagger\,(\delta\mathbf{A})\,\underline{m}_i$$

car $\delta(\underline{m}_i^\dagger\,\underline{m}_i = 1) = 0$. 2) Pour $\delta\mathbf{A} = 0$ cette relation entraîne $\delta(\underline{m}_i^\dagger\mathbf{A}\underline{m}_i) = 0$; donc tout vecteur propre normé extrémalise $(\underline{m}^\dagger\mathbf{A}\underline{m})$. Cette propriété est utilisée pour la recherche des vecteurs et valeurs propres par des méthodes variationnelles (**Rayleigh-Ritz**).

Oscillateurs couplés. Pour $\mathbf{F}$ "petite", on trouve les pulsations propres de

$$\mathbf{M}\ddot{\underline{x}} + \mathbf{F}\dot{\underline{x}} + \mathbf{K}\underline{x} = 0$$

(*cf.* section 4.1.3) en posant $\underline{x}(t) = e^{i(\omega+\delta\omega)t}(\underline{m} + \delta\underline{m})$. A l'ordre le plus bas il vient $(\mathbf{K} - \omega^2\mathbf{M})\delta\underline{m} + (-2\omega\,\delta\omega\mathbf{M} + i\omega\mathbf{F})\,\underline{m} = 0$, et en multipliant à gauche par $\underline{m}^t$ on obtient $\delta\omega = \frac{i}{2}\frac{\underline{m}^t\mathbf{F}\underline{m}}{\underline{m}^t\mathbf{M}\underline{m}}$ (imaginaire pur en raison des pertes d'énergie). 2) On montre de même pour l'équation $(\mathbf{M} + \delta\mathbf{M})\,\ddot{\underline{x}} + (\mathbf{K} + \delta\mathbf{K})\,\underline{x} = 0$ que $\delta\omega = \frac{\underline{m}^t(\delta\mathbf{K} - \omega^2\delta\mathbf{M})\underline{m}}{2\omega\underline{m}^t\mathbf{M}\underline{m}}$.

▷ Fonctions propres

Les notions de valeurs et vecteurs propres s'étendent aux applications (opérateurs) linéaires sur les fonctions. Par exemple les translations de temps $(T_\tau\,f)(t) = f(t-\tau)$ pour les fonctions de période T ont les exponentielles $e^{-i\omega t}$ ($\omega = \dfrac{2\pi n}{T}$) comme fonctions propres avec $e^{i\omega\tau}$ comme valeurs propres. Plus généralement des champs $f(\vec{r},t)$ ayant une dépendance de type onde plane en $e^{-i(\omega t - \vec{k}\cdot\vec{r})}$ sont, au moins formellement, fonctions propres des opérations $T_{\vec{a},\tau}$ de translation spatiale et temporelle, définies par $T_{\vec{a},\tau}\,f(\vec{r},t) = f(\vec{r}-\vec{a},t-\tau)$ avec les valeurs propres $e^{i\omega\tau}\,e^{-i\vec{k}\cdot\vec{a}}$ (résultat utilisé à la section 3.4.2).

AUTRES EXEMPLES. La **symétrisation** ou l'**antisymétrisation** d'une fonction de n variables, qui intervient quand on construit les fonctions d'onde de bosons ou de fermions, correspond aux opérations linéaires

$$f(x_1,\cdots,x_n) \to (Sf)(x_1,\cdots,x_n) = (n!)^{-1} \sum_P f(x_{P(1)},\cdots,x_{P(n)})$$
$$\text{ou} \quad f(x_1,\cdots,x_n) \to (Af)(x_1,\cdots,x_n) = (n!)^{-1} \sum_P (-1)^{|P|} f(x_{P(1)},\cdots,x_{P(n)}) \ .$$

$\sum_P$ est la somme sur toutes les permutations, au nombre de $n!$. Par exemple pour $n=2$:

$$f_1(x_1)\, f_2(x_2) \to \tfrac{1}{2}\,[f_1(x_1)\, f_2(x_2) \pm f_1(x_2)\, f_2(x_1)].$$

Si $\mathcal{P}$ est l'opération de permutation définie par $(\mathcal{P}f)(x_{P(1)},\cdots,x_{P(n)}) = f(x_1,\cdots,x_n)$, on vérifie facilement que Sf et Af, qui sont respectivement symétrique et antisymétrique, sont fonctions propres de $\mathcal{P}$ avec les valeurs propres 1 et $(-1)^{|P|}$.

4.2.4 Matrices de Pauli ; groupe de symétrie de la physique ; spineurs

Ces matrices 2×2, forme moderne des quaternions inventés par Hamilton pour généraliser les nombres complexes et pour algébriser l'espace à trois dimensions, sont utiles non seulement pour obtenir la loi de composition des rotations, mais aussi pour décrire l'espace-temps en relativité, et traiter les systèmes à deux niveaux en quantique (*cf.* section 4.4.2).

▷ Définition et propriétés

Les **matrices de Pauli** (hermitiennes de trace nulle)

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

forment avec l'identité $\mathbf{I}$ une base des matrices 2×2. La matrice la plus générale peut en effet s'écrire :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} V_0+V_z & V_x-iV_y \\ V_x+iV_y & V_0-V_z \end{pmatrix} = V_0\,\mathbf{I} + \overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{V} = \mathbf{V} \ .$$

(La notation vectorielle $\overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{V} = V_x\,\boldsymbol{\sigma}_x + V_y\,\boldsymbol{\sigma}_y + V_z\,\boldsymbol{\sigma}_z$ est justifiée par ce qui suit.) $\mathbf{V}$ est hermitienne si V_0 et $\overrightarrow{V}$ sont réels. Dans ce cas on peut poser $\overrightarrow{V} = V\,\hat{n}$ avec $\hat{n} \equiv (\theta,\varphi)$:

$$\mathbf{V} = V_0\,\mathbf{I} + V\,\overrightarrow{\boldsymbol{\sigma}}\cdot\hat{n} = V_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + V \begin{pmatrix} \cos\theta & \sin\theta e^{-i\varphi} \\ \sin\theta e^{i\varphi} & -\cos\theta \end{pmatrix} \ .$$

Les **valeurs propres** et **vecteurs propres** normés de $\mathbf{V}$ sont :

$$\lambda_\pm = V_0 \pm V \quad , \quad \underline{m}_+ = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}\, e^{i\varphi} \end{pmatrix} \quad , \quad \underline{m}_- = \begin{pmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2}\, e^{i\varphi} \end{pmatrix} \quad , \quad (\overrightarrow{\boldsymbol{\sigma}}\cdot\hat{n})\,\underline{m}_\pm = \pm\underline{m}_\pm \ .$$

DÉMONSTRATION. On considère le cas $V_0 = 0$, ce qui revient simplement à décaler les valeurs propres (de $-V_0$) sans changer les vecteurs propres. $\det(\lambda\mathbf{I}-\mathbf{H}) = \lambda^2 - V^2 = 0$ donne $\lambda_\pm$. Les composantes x et y des vecteurs propres se déduisent de $\cos\theta\, x + \sin\theta\, e^{-i\varphi}\, y = \pm x$ (et de la normalisation $|x|^2+|y|^2=1$) ; ces vecteurs sont définis à un facteur $e^{i\phi}$ près. Les **projecteurs** associés sont : $\mathbf{P}_\pm = (\mathbf{I} \pm \overrightarrow{\boldsymbol{\sigma}}\cdot\hat{n})/2$.

Les matrices $\overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{V}$ vérifient la loi de multiplication (conséquence de $\boldsymbol{\sigma}_x^2 = \boldsymbol{\sigma}_y^2 = \boldsymbol{\sigma}_z^2 = \mathbf{I}$ et de $\boldsymbol{\sigma}_x\,\boldsymbol{\sigma}_y = -\boldsymbol{\sigma}_y\,\boldsymbol{\sigma}_x = i\boldsymbol{\sigma}_z \cdots$) :

$$\big(\vec{\boldsymbol{\sigma}}\cdot\overrightarrow{V}_1\big)\big(\vec{\boldsymbol{\sigma}}\cdot\overrightarrow{V}_2\big)=(\overrightarrow{V}_1\cdot\overrightarrow{V}_2)\,\mathbf{I}+i\vec{\boldsymbol{\sigma}}\cdot(\overrightarrow{V}_1\wedge\overrightarrow{V}_2)\ .$$

Elle permet d'écrire (car $(\vec{\boldsymbol{\sigma}}\cdot\hat{n})^2=\mathbf{I}$)) les exponentielles

$$\mathbf{U}_{\hat{n}}(\theta)=e^{-\frac{i}{2}\,\vec{\boldsymbol{\sigma}}\cdot\hat{n}\,\theta}=\cos\frac{\theta}{2}\,\mathbf{I}-i\sin\frac{\theta}{2}\,\vec{\boldsymbol{\sigma}}\cdot\hat{n}\quad\text{(matrices unitaires)}\ ,$$

$$\mathbf{H}_{\hat{n}}(\varphi)=e^{-\frac{1}{2}\,\vec{\boldsymbol{\sigma}}\cdot\hat{n}\,\varphi}=\cosh\frac{\varphi}{2}\,\mathbf{I}-\sinh\frac{\varphi}{2}\,\vec{\boldsymbol{\sigma}}\cdot\hat{n}\quad\text{(matrices hermitiennes)}\ ,$$

qui, on va le voir, constituent une **représentation** des rotations et des transformations de Lorentz pures. En particulier :

$$\mathbf{U}_{\hat{x}}(\theta)=\begin{pmatrix}\cos\frac{\theta}{2} & -i\sin\frac{\theta}{2}\\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2}\end{pmatrix}\quad;\quad\mathbf{U}_{\hat{z}}(\theta)=\begin{pmatrix}e^{-i\frac{\theta}{2}} & 0\\ 0 & e^{i\frac{\theta}{2}}\end{pmatrix}\ .$$

▷ Matrices de Pauli et espace-temps

A un intervalle d'espace-temps $(cT,\ \overrightarrow{R})$ de coordonnées (cT,X,Y,Z) (*cf.* section 3.1.2) on peut associer mathématiquement la matrice hermitienne

$$\mathbf{X}=cT\,\mathbf{I}+\overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{R}\quad,\quad\det\mathbf{X}=c^2T^2-\overrightarrow{R}^2\ .$$

Les transformations, avec $\mathbf{M}$ matrices 2×2 complexes de déterminant 1,

$$\mathbf{X}'=\mathbf{M}\mathbf{X}\mathbf{M}^\dagger\qquad(\det\mathbf{M}=1)$$

induisent sur (cT,X,Y,Z) des changements qui définissent le **groupe de Lorentz** : transformations linéaires telles que $c^2T^2-\overrightarrow{R}^2=c^2T'^2-\overrightarrow{R}'^2$. En effet, elles conservent le caractère hermitien de $\mathbf{X}$, ce qui permet d'écrire $\mathbf{X}'=cT'\,\mathbf{I}+\overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{R}'$, et son déterminant. On vérifie aussi que les matrices $\mathbf{M}$ (4 éléments complexes et déterminant réel égal à 1) dépendent de $8-2=6$ paramètres réels comme ce groupe.

Rotations. Si $\mathbf{M}$ est unitaire ($\mathbf{M}^\dagger=\mathbf{M}^{-1}$) la trace $2cT$ de $\mathbf{X}$ est inchangée donc aussi $|\overrightarrow{R}|$. En considérant la transformation infinitésimale de $\overrightarrow{R}$ associée à une matrice $\mathbf{U}_{\hat{n}}(\epsilon)$

$$\vec{\boldsymbol{\sigma}}\cdot\overrightarrow{R}'=\big(\mathbf{I}-i\vec{\boldsymbol{\sigma}}\cdot\hat{n}\,\frac{\epsilon}{2}\big)\vec{\boldsymbol{\sigma}}\cdot\overrightarrow{R}\big(\mathbf{I}+i\vec{\boldsymbol{\sigma}}\cdot\hat{n}\,\frac{\epsilon}{2}\big)\ ,$$

on obtient $\overrightarrow{R}'=\overrightarrow{R}+\epsilon\hat{n}\wedge\overrightarrow{R}$ et on en déduit que $\mathbf{U}_{\hat{n}}(\theta)$ correspond à une rotation d'axe $\hat{n}$ et d'angle θ. (On peut aussi faire le calcul pour θ fini et retrouver le résultat de la section 3.3.1.) On notera que $\mathbf{U}_{\hat{n}}(\theta)$ et $-\mathbf{U}_{\hat{n}}(\theta)=\mathbf{U}_{\hat{n}}(\theta+2\pi)$ correspondent à la même rotation de $\overrightarrow{R}$. La loi de multiplication des $\mathbf{U}_{\hat{n}}(\theta)$

$$\mathbf{U}_{\hat{n}_2}(\theta_2)\mathbf{U}_{\hat{n}_1}(\theta_1)=\pm\mathbf{U}_{\hat{n}}(\theta)$$

s'écrit, en posant $c=\cos\dfrac{\theta}{2}$ et $\vec{s}=\sin\dfrac{\theta}{2}\,\hat{n}$ (petit calcul) :

$$c=c_1c_2-\vec{s}_1\cdot\vec{s}_2\ \ \text{et}\ \ \vec{s}=c_2\vec{s}_1+c_1\vec{s}_2+\vec{s}_2\wedge\vec{s}_1\ ;$$

modulo l'équivalence $(c,\vec{s})\equiv(-c,-\vec{s})$, c'est la **loi de composition des rotations** (formule de Hamilton). On retrouve que seules les rotations de même axe ($\vec{s}_1\parallel\vec{s}_2$), ou d'angle π et d'axes perpendiculaires ($c_1=c_2=0$ et $\vec{s}_1\perp\vec{s}_2$) commutent (*cf.* section 3.1.3).

REMARQUE. Les matrices $\mathbf{U}_{\hat{n}}(\theta)$ décrivent en physique l'action des rotations sur les états de particules de spin $\frac{1}{2}$ au repos (électrons, protons, neutrons, quarks), et la propriété étonnante $\mathbf{U}_{\hat{n}}(2\pi) = -\mathbf{I}$ (et non $\mathbf{I}$) a été vérifiée expérimentalement. Elle signifie que le "vrai" groupe de symétrie n'est pas le groupe des rotations des vecteurs de $\mathbb{R}^3$, mais le groupe SU(2) des matrices 2×2 unitaires "spéciales" (c.a.d. de déterminant 1).

Transformations de Lorentz pures. On vérifie de même que la relation

$$cT'\mathbf{I} + \vec{\boldsymbol{\sigma}} \cdot \overrightarrow{R}' = (\mathbf{I} - \vec{\boldsymbol{\sigma}} \cdot \hat{n}\, \frac{\epsilon}{2})(cT\mathbf{I} + \vec{\boldsymbol{\sigma}} \cdot \overrightarrow{R})(\mathbf{I} - \vec{\boldsymbol{\sigma}} \cdot \hat{n}\, \frac{\epsilon}{2}) \ ,$$

correspond (*cf.* section 4.3.3) à une transformation (passive) de Lorentz pure infinitésimale

$$cT' = cT - \overrightarrow{\epsilon} \cdot \overrightarrow{R} \quad , \quad \overrightarrow{R}' = \overrightarrow{R} - \overrightarrow{\epsilon}\, cT \ .$$

REMARQUE. On peut aussi associer à $(cT, \overrightarrow{R})$ la matrice $\widetilde{\mathbf{X}}$ **conjuguée** de $\mathbf{X}$:

$$\widetilde{\mathbf{X}} = cT\,\mathbf{I} - \overrightarrow{\boldsymbol{\sigma}} \cdot \overrightarrow{R} = \boldsymbol{\epsilon}\,\overline{\mathbf{X}}\,\boldsymbol{\epsilon}^{-1} \quad \text{avec} \quad \boldsymbol{\epsilon} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ .$$

(On vérifie que pour chaque matrice de Pauli : $\boldsymbol{\epsilon}\overline{\boldsymbol{\sigma}_i}\boldsymbol{\epsilon}^{-1} = -\boldsymbol{\sigma}_i$, $i = x, y, z$.) Alors les transformations $\mathbf{X} \to \mathbf{X}'$ équivalent à :

$$\widetilde{\mathbf{X}}' = \widetilde{\mathbf{M}}\widetilde{\mathbf{X}}\widetilde{\mathbf{M}}^{\dagger} \quad \text{avec} \quad \widetilde{\mathbf{M}} = \boldsymbol{\epsilon}\overline{\mathbf{M}}\boldsymbol{\epsilon}^{-1} = \mathbf{M}^{\dagger -1} \ (\text{car } \det \mathbf{M} = \mathbf{1}) \ .$$

Les matrices $\widetilde{\mathbf{M}}$ constituent une seconde représentation à deux dimensions du groupe de Lorentz. Elle ne peut être ramenée à la première par un changement de base car il n'existe pas de matrice $\mathbf{P}$ telle que $\widetilde{\mathbf{M}} = \mathbf{PMP}^{-1}$ pour tout $\mathbf{M}$ (celà impliquerait $\mathbf{U} = \mathbf{PUP}^{-1}$ pour toute matrice unitaire).

▷ Spineurs

On appelle ainsi des vecteurs complexes à deux composantes qui se transforment selon :

$$\underline{\psi}\,' = \mathbf{M}\,\underline{\psi} \quad \text{ou} \quad \underline{\widetilde{\psi}}\,' = \widetilde{\mathbf{M}}\,\underline{\widetilde{\psi}} \ .$$

($\underline{\psi}$ et $\underline{\widetilde{\psi}}$ se transforment avec la même loi pour une rotation car alors $\mathbf{M} = \widetilde{\mathbf{M}}$). On vérifie aisément que $\boldsymbol{\epsilon}\underline{\overline{\psi}}$ se transforme comme $\underline{\widetilde{\psi}}$ (et $\boldsymbol{\epsilon}\underline{\overline{\widetilde{\psi}}}$ comme $\underline{\psi}$). Par exemple : $\boldsymbol{\epsilon}\underline{\overline{\psi}}\,' = \boldsymbol{\epsilon}\overline{\mathbf{M}}(\boldsymbol{\epsilon}^{-1}\boldsymbol{\epsilon})\underline{\overline{\psi}} = \widetilde{\mathbf{M}}(\boldsymbol{\epsilon}\underline{\overline{\psi}})$. $\underline{\psi}$ et $\underline{\widetilde{\psi}}$ sont dits **spineurs conjugués** (application section 4.4.5).

4.2.5 Groupe de rotation et classification des grandeurs physiques

On a vu au chapitre 3, avec le principe de Curie, l' importance des lois de transformation des grandeurs (et champs) physiques. L'objet de cette section est de montrer que pour les rotations ces grandeurs ne se limitent pas aux scalaires et aux vecteurs.

▷ Classification des matrices 3×3 réelles

Soit $\overrightarrow{W} = \mathbf{M}\,\overrightarrow{V}$ une relation entre grandeurs vectorielles réelles relatives à un système physique. Pour le système tourné les grandeurs et la relation deviennent $\overrightarrow{W}' = \mathbf{R}\,\overrightarrow{W}$, $\overrightarrow{V}' = \mathbf{R}\,\overrightarrow{V}$ et $\overrightarrow{W}' = \mathbf{M}'\,\overrightarrow{V}\,'$ avec $\mathbf{M}' = \mathbf{RMR}^t$. Si $\mathbf{M} = \lambda\mathbf{I}$, ou $\mathbf{M} = \mathbf{A}$ antisymétrique, ou $\mathbf{M} = \widetilde{\mathbf{M}}$ symétrique de trace nulle, $\mathbf{M}'$ a les mêmes propriétés que $\mathbf{M}$. On en déduit que pour une matrice quelconque la décomposition

$$\mathbf{M} = \frac{1}{3}\,(\mathrm{tr}\mathbf{M})\,\mathbf{I} + \frac{1}{2}(\mathbf{M} - \mathbf{M}^t) + \widetilde{\mathbf{M}} \quad (\text{où } \widetilde{M}_{ij} = \frac{1}{2}\,(M_{ij} + M_{ji}) - \frac{1}{3}\,\delta_{ij} \sum_k M_{kk})$$

est invariante par rotation. La matrice $\lambda\mathbf{I}$ invariante est une **grandeur scalaire** ; l'action d'une matrice antisymétrique sur $\overrightarrow{V}$ pouvant toujours s'écrire sous la forme d'un produit vectoriel (*cf.* exemple 2 section 4.2.2) la matrice $\mathbf{A} = \frac{1}{2}(\mathbf{M} - \mathbf{M}^t)$ est assimilable à une **grandeur vectorielle** ; $\mathbf{Q} = \widetilde{\mathbf{M}}$ caractérise une **grandeur quadrupolaire**. L'E.V. des matrices 3×3 est ainsi décomposé en une somme directe de trois sous espaces laissés invariants par les rotations ; pour les dimensions, on vérifie que $3 \times 3 = 1 + 3 + 5$ ($\mathbf{A}$ dépend de trois paramètres et $\mathbf{Q}$ de cinq).

EXEMPLES. Le **développement multipolaire** au second ordre du potentiel électrostatique à grande distance créé par une distribution de charges est (*cf.* $|\vec{r} - \vec{r}_\alpha|^{-1} \simeq r^{-1}\left(1 - 2\dfrac{\vec{r}_\alpha \cdot \hat{r}}{r} + \dfrac{r_\alpha^2}{r^2}\right)^{-\frac{1}{2}}$ et section 5.1.2)

$$V(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_\alpha \frac{q_\alpha}{|\vec{r} - \vec{r}_\alpha|} = \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r} + \frac{\vec{p}\cdot\hat{r}}{r^2} + \frac{\hat{r}\cdot\mathbf{Q}\hat{r}}{2r^3} + \cdots\right) ,$$

avec $q = \sum_\alpha q_\alpha$, $\vec{p} = \sum_\alpha q_\alpha \vec{r}_\alpha$ et $Q_{ij} = \sum_\alpha q_\alpha \left(3x_{i\alpha}\, x_{j\alpha} - r_\alpha^2\, \delta_{ij}\right)$. Pour le potentiel de gravitation, en prenant l'origine au centre de masse, on obtient de même

$$V(\vec{r}) = -G \sum_\alpha m_\alpha\, |\vec{r} - \vec{r}_\alpha|^{-1} = -G\left(r^{-1}\, m + (2r^3)^{-1}\, \hat{r}\cdot\mathbf{Q}\hat{r} + \cdots\right)$$

avec $m = \sum_\alpha m_\alpha$ et $Q_{ij} = \sum_\alpha m_\alpha \left(3x_{i\alpha}\, x_{j\alpha} - r_\alpha^2\, \delta_{ij}\right)$ (pas de terme dipolaire).

▷ Représentations linéaires du groupe des rotations

Une action linéaire (*cf.* section 3.1.1) du groupe des rotations (en fait du groupe SU(2)) sur les paramètres $\underline{x} \equiv (x_1, \cdots, x_n)$ qui caractérisent un système s'écrit $\underline{x}' = \mathcal{R}(\mathbf{U})\,\underline{x}$ avec $\mathcal{R}(\mathbf{U}_1\,\mathbf{U}_2) = \mathcal{R}(\mathbf{U}_1)\,\mathcal{R}(\mathbf{U}_2)$. $\mathcal{R}(\mathbf{U})$ est une matrice $n \times n$ qui représente la rotation dans l'espace vectoriel des paramètres $X = \{\underline{x}\}$ (matrice orthogonale 3×3 lorsqu'elle agit sur les vecteurs de l'espace ordinaire). La représentation est dite **irréductible** si il n'existe pas de sous espace de X laissé globalement invariant.

On démontre (*cf.* ouvrages de physique quantique) que pour toute dimension n, écrite traditionnellement

$$n = 2j + 1 \quad \text{avec } j \text{ demi-entier positif ou nul,}$$

il existe une représentation irréductible unique, à un changement de base près. Les grandeurs **scalaires** (invariantes), **vectorielles**, **quadrupolaires** et les **spineurs** ci-dessus se transforment selon les représentations $j = 0, 1, 2$ ($n = 1, 3, 5$) et $j = \frac{1}{2}$ ($n = 2$, matrices $\mathbf{U}_{\hat{n}}(\theta)$). Pour j fixé quelconque, la base

$$\{\underline{e}_{jm}\} \quad (m = -j, -j+1, \cdots, j)$$

peut être choisie de sorte que les matrices soient unitaires, et que les $\underline{e}_{jm}$ soient vecteurs propres des rotations d'angle α autour de $\hat{z}$ avec les valeurs propres $e^{-im\alpha}$. Pour $j = \frac{1}{2}$ les vecteurs $\underline{e}_{\frac{1}{2}\frac{1}{2}}$ et $\underline{e}_{\frac{1}{2}-\frac{1}{2}}$, notés aussi ci-dessous $\underline{\psi}_+$ et $\underline{\psi}_-$, correspondent à $\begin{pmatrix}1\\0\end{pmatrix}$ et $\begin{pmatrix}0\\1\end{pmatrix}$. Pour $j = 1$ les vecteurs $\underline{e}_{11}$, $\underline{e}_{10}$ et $\underline{e}_{1-1}$ sont proportionnels à $\hat{x} + i\hat{y}$, $\hat{z}$ et $\hat{x} - i\hat{y}$. Des exemples d'espaces de représentation avec $j = l$ entier sont les E.V. de fonctions engendrés par la base des harmoniques sphériques $Y_l^m(\theta,\varphi) \equiv \underline{e}_{lm}$. On vérifie que $Y_l^m(\theta, \varphi - \alpha) = e^{-im\alpha}\, Y_l^m(\theta,\varphi)$ (*cf.* section 4.1.2). Les **orbitales atomiques** $s, p, d, f \cdots$ sont donc associées aux représentations $l = 0, 1, 2, 3, \cdots$.

On montre aussi que l'E.V. produit tensoriel $E_{j_1} \otimes E_{j_2}$ (de dimension $(2j_1 + 1)(2j_2 + 1)$) engendré par les vecteurs $\underline{e}_{j_1 m_1} \otimes \underline{e}_{j_2 m_2}$, qui sont chacun changé en $\mathcal{R}_1\,\underline{e}_{j_1 m_1} \otimes \mathcal{R}_2\,\underline{e}_{j_2 m_2}$ dans une rotation, se décompose en une somme directe $\oplus_j E_j$ de sous E.V. invariants, la somme sur j portant sur les valeurs

$$j = |j_1 - j_2|, |j_1 - j_2 + 1|, \cdots, j_1 + j_2 \ .$$

Cette propriété correspond en physique quantique à la **composition des moments cinétiques**. Par exemple si $j_1 = j_2 = \frac{1}{2}$ (E.V. engendré par les quatre vecteurs $\underline{\psi}_\pm \otimes \underline{\psi}_\pm$) alors $j = 0$ ou 1. On vérifie facilement que $2^{-\frac{1}{2}}\,(\underline{\psi}_+ \otimes \underline{\psi}_- - \underline{\psi}_- \otimes \underline{\psi}_+)$ est invariant (conséquence de $\underline{\psi}_+ \to U_{11}\underline{\psi}_+ + U_{21}\underline{\psi}_-$, $\underline{\psi}_- \to U_{12}\underline{\psi}_+ + U_{22}\underline{\psi}_-$ et de $\det \mathbf{U} = U_{11}U_{22} - U_{12}U_{21} = 1$). L'E.V. complémentaire $j = 1$ est engendré par $\underline{\psi}_+ \otimes \underline{\psi}_+$, $2^{-\frac{1}{2}}\,(\underline{\psi}_+ \otimes \underline{\psi}_- + \underline{\psi}_- \otimes \underline{\psi}_+)$ et $\underline{\psi}_- \otimes \underline{\psi}_-$. En physique classique la décomposition des matrices 3×3 en scalaire, vecteur et quadrupole correspond au cas $j_1 = j_2 = 1$.

4.3 APPLICATIONS EN PHYSIQUE CLASSIQUE

4.3.1 Déformations et contraintes ; élasticité ; viscosité

Un corps se déforme lorsque les distances entre ses points varient. On se limite ici aux déformations infinitésimales et aux lois linéaires entre déformations et contraintes.

▷ Déformations linéaires infinitésimales

Elles correspondent à des transformations qui changent les vecteurs $\vec{l} = \overrightarrow{AB}$ en $\vec{l}\,' = \vec{l} + \overrightarrow{\delta l}$ avec $\overrightarrow{\delta l} = \mathbf{M}\,\vec{l}$, $\mathbf{M}$ étant une matrice "infinitésimale". Leur composition revient à multiplier des matrices de la forme $\mathbf{I} + \mathbf{M}$, c.a.d. à ajouter les matrices infinitésimales $\mathbf{M}$, car $(\mathbf{I} + \mathbf{M_1})\,(\mathbf{I} + \mathbf{M_2}) = \mathbf{I} + \mathbf{M_1} + \mathbf{M_2}$ à l'ordre le plus bas ; ceci équivaut à ajouter les $\overrightarrow{\delta l}$. Comme une rotation infinitésimale, qui correspond à une matrice $\mathbf{M}$ antisymétrique, ne change pas les distances, les déformations sont caractérisées par les matrices symétriques notées $[\boldsymbol{\epsilon}]$:

$$\begin{pmatrix} \delta l_x \\ \delta l_y \\ \delta l_z \end{pmatrix} = \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{xy} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{xz} & \epsilon_{yz} & \epsilon_{zz} \end{pmatrix} \begin{pmatrix} l_x \\ l_y \\ l_z \end{pmatrix} \ ;$$

ces **matrices de déformation** dépendent de six paramètres.

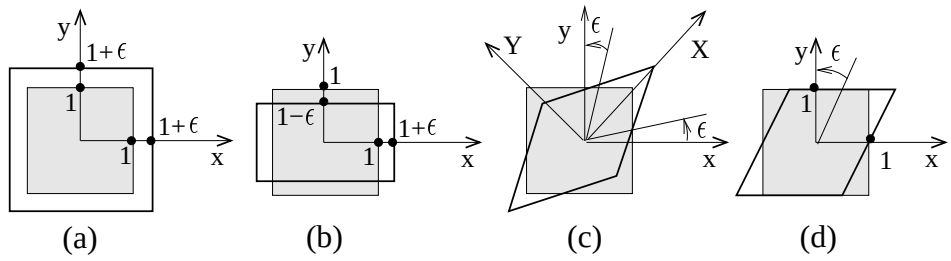
EXEMPLES. Les translations ne changeant pas les distances entre points, on considère les déformations qui laissent l'origine O fixe : $\vec{r}\,' = \vec{r} + [\boldsymbol{\epsilon}]\vec{r}$. Les figures 5a,b,c montrent l'effet sur le carré $[-1, 1] \times [-1, 1]$ des trois transformations planes

$$(a) \qquad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1+\epsilon_1 & 0 \\ 0 & 1+\epsilon_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \ ,$$

$$(b) \qquad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1+\epsilon_2 & 0 \\ 0 & 1-\epsilon_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \ ,$$

$$(c) \qquad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & \epsilon_3 \\ \epsilon_3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \ ,$$

dont toute déformation infinitésimale plane est la composée. (a) est une **dilatation** (homothétie) de rapport $1+\epsilon_1$. (b) et (c), qui ne changent pas la surface, sont appelées **cisaillement** ; (c) se ramène à (b), pour $\epsilon_3 = \epsilon_2$, par une rotation de $\dfrac{\pi}{4}$ des axes.



FIGURES 5

Enfin la figure 5d montre l'effet de la transformation $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & \epsilon \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ appelée **cisaillement simple** ; elle résulte de la composition d'un cisaillement "pur" de type (c) avec $\epsilon_3 = \dfrac{\epsilon}{2}$, et d'une rotation d'angle $-\dfrac{\epsilon}{2}$, comme le montre les figures 5c et 5d ou l'addition des matrices $\begin{pmatrix} 0 & \frac{\epsilon}{2} \\ \frac{\epsilon}{2} & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & \frac{\epsilon}{2} \\ -\frac{\epsilon}{2} & 0 \end{pmatrix}$.
En considérant la déformation des vecteurs $\hat{x}$, $\hat{y}$ et $\hat{z}$, par exemple $\hat{x} \to (1+\epsilon_{xx})\,\hat{x} + \epsilon_{xy}\,\hat{y} + \epsilon_{xz}\,\hat{z}$, on voit que ϵ_{xx} est l'**allongement relatif d'un vecteur parallèle à** $\hat{x}$, que $2\epsilon_{xy}$ est le **rapprochement angulaire** (symétrique) des directions x et y, et que le volume du cube construit sur $(\hat{x},\hat{y},\hat{z})$ devient $\det(\mathbf{I}+[\boldsymbol{\epsilon}]) = 1+\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}$; $\mathrm{tr}[\boldsymbol{\epsilon}]$ est donc la **variation relative de volume**. La relation $[\boldsymbol{\epsilon}] = \frac{1}{3}\,\mathrm{tr}[\boldsymbol{\epsilon}]\,\mathbf{I} + [\widetilde{\boldsymbol{\epsilon}}]$, ou $\epsilon_{ij} = \frac{1}{3}\sum_k \epsilon_{kk}\,\delta_{ij} + (\epsilon_{ij} - \frac{1}{3}\sum_k \epsilon_{kk}\,\delta_{ij})$, montre que toute déformation se décompose en une dilatation (homothétie), associée à la variation de volume, et des cisaillements (qui ne modifient pas le volume car $\mathrm{tr}[\widetilde{\boldsymbol{\epsilon}}] = 0$). Enfin dans la base $\widehat{X},\widehat{Y},\widehat{Z}$ qui rend $[\boldsymbol{\epsilon}]$ diagonale, la déformation apparaît comme la composition d'affinités $1+\epsilon_{_X}$, $1+\epsilon_{_Y}$, et $1+\epsilon_{_Z}$ dans les directions $\widehat{X},\widehat{Y},\widehat{Z}$ ($\epsilon_{_X}$, $\epsilon_{_Y}$, $\epsilon_{_Z}$ valeurs propres de $[\,\boldsymbol{\epsilon}\,]$).

Champ de déformations. Si chaque point $\vec{r}$ se déplace de $\overrightarrow{\psi}(\vec{r},t)$, le vecteur $\overrightarrow{\mathrm{d}r}$ joignant deux points voisins devient $\overrightarrow{\mathrm{d}r} + \overrightarrow{\mathrm{d}\psi}$ avec par exemple (*cf.* section 7.1.1) $\mathrm{d}\psi_x = \psi_x(\vec{r}+\overrightarrow{\mathrm{d}r},t) - \psi_x(\vec{r},t) = \partial_x\psi_x\,\mathrm{d}x + \partial_y\psi_x\,\mathrm{d}y + \partial_z\psi_x\,\mathrm{d}z$, ou plus généralement $\mathrm{d}\psi_i = \sum_j (\partial_j\psi_i)\,\mathrm{d}x_j$ soit $\mathrm{d}\overrightarrow{\psi} = \mathbf{M}\,\overrightarrow{\mathrm{d}r}$. Si $\mathbf{M}$ est infinitésimale, sa partie antisymétrique traduit une rotation de $\overrightarrow{\mathrm{d}r}$; sa partie symétrique $[\boldsymbol{\epsilon}]$ correspond à une déformation dont la matrice a pour éléments et trace

$$\epsilon_{ij} = \tfrac{1}{2}(\partial_i\psi_j + \partial_j\psi_i) \text{ et } \mathrm{tr}[\boldsymbol{\epsilon}] = \mathrm{div}\,\overrightarrow{\psi} = \tfrac{\delta V}{V}\ .$$

$\frac{\delta V}{V}$représente une **variation relative locale de volume**. Si le déplacement $\overrightarrow{\psi}$ s'effectue dans le temps $\mathrm{d}t$ on écrit $\overrightarrow{\psi} = \vec{v}\,\mathrm{d}t$; la **vitesse de déformation** est définie par la matrice notée $[\dot{\boldsymbol{\epsilon}}]$, d'éléments $[\dot{\boldsymbol{\epsilon}}]_{ij} = \frac{1}{2}\,(\partial_i v_j + \partial_j v_i)$ et de trace $\mathrm{tr}[\dot{\boldsymbol{\epsilon}}] = \mathrm{div}\,\overrightarrow{v} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta V}{V}$ (*cf.* section 7.1.3).

▷ Contraintes

Soient, au sein de la matière, un élément de surface $\overrightarrow{\mathrm{d}S}$ orienté de 1 vers 2 et $\overrightarrow{\mathrm{d}F} \equiv \overrightarrow{\mathrm{d}F}_{2\to1}$ la force de contact exercée par la région 2 sur la région 1. La relation de proportionnalité $\overrightarrow{\mathrm{d}F} = [\boldsymbol{\sigma}]\,\overrightarrow{\mathrm{d}S} = -[\boldsymbol{p}]\,\overrightarrow{\mathrm{d}S}$ (donc $[\boldsymbol{p}]\,\overrightarrow{\mathrm{d}S} = \overrightarrow{\mathrm{d}F}_{1\to2}$) définit les **matrices de contrainte** $[\boldsymbol{\sigma}]$ **et de pression** $[\boldsymbol{p}]$. Explicitement :

$$\begin{pmatrix} \mathrm{d}F_x \\ \mathrm{d}F_y \\ \mathrm{d}F_z \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \begin{pmatrix} \mathrm{d}S_x \\ \mathrm{d}S_y \\ \mathrm{d}S_z \end{pmatrix} \qquad \text{avec } \sigma_{xy} = \sigma_{yx},\ \sigma_{yz} = \sigma_{zy},\ \sigma_{zx} = \sigma_{xz}\ .$$

En prenant $\overrightarrow{\mathrm{d}S} = \mathrm{d}S\,\hat{x}$, on voit que σ_{xx} et $(\sigma_{xy}, \sigma_{xz})$ représentent les composantes normale et tangentielles de la force exercée sur une surface unité orientée selon $\hat{x}$. La figure 6 représente, en perspective et dans le cas où les σ_{ij} sont positifs et constants, les forces exercées sur les faces d'un cube de côté unité. Le caractère symétrique de $[\boldsymbol{\sigma}]$ admis ici est discuté à la section 7.2.3. Comme pour les déformations on pose $[\boldsymbol{\sigma}] = [\boldsymbol{\sigma}]_{\text{isotrope}} + [\widetilde{\boldsymbol{\sigma}}]$ (avec $[\boldsymbol{\sigma}]_{\text{isotrope}} = \dfrac{1}{3}\,\mathrm{tr}[\boldsymbol{\sigma}]\,\mathbf{I}$ et $\mathrm{tr}[\widetilde{\boldsymbol{\sigma}}] = 0$). Si $[\widetilde{\boldsymbol{\sigma}}] = 0$ alors $[\boldsymbol{\sigma}] = [\boldsymbol{\sigma}]_{\text{isotrope}} = -p\,\mathbf{I}$ et on retrouve l'expression $\overrightarrow{\mathrm{d}F} = -p\,\overrightarrow{\mathrm{d}S}$ de la force de pression due à un fluide parfait.

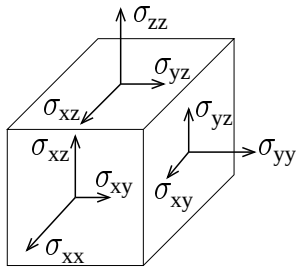


FIGURE 6

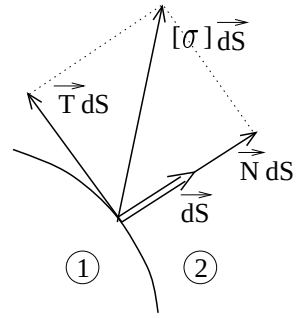


FIGURE 7

REMARQUE. La figure 7 représente les composantes normale N et tangentielle T de la force par unité de surface. Dans la base où $[\boldsymbol{\sigma}]$ est diagonale, et où les composantes du vecteur unitaire $\hat{n}$ de $\overrightarrow{\mathrm{d}S}$ sont (α, β, γ), on a : $N = \sigma_X\,\alpha^2 + \sigma_Y\,\beta^2 + \sigma_Z\,\gamma^2$ et $T^2 + N^2 = \sigma_X^2\,\alpha^2 + \sigma_Y^2\,\beta^2 + \sigma_Z^2\,\gamma^2$. Ces relations permettent d'illustrer comment N et T dépendent de la direction $\hat{n}$. Le domaine de variation du couple (N, T) en fonction de (α, β, γ) (en gris sur la figure 8 dans le cas $\sigma_X > \sigma_Y > \sigma_Z$) est limité par les **cercles de Mohr** ; par exemple si $\gamma = 0$, en éliminant α et β on obtient le cercle $T^2 + \big(N - \frac{\sigma_X + \sigma_Y}{2}\big)^2 = \big(\frac{\sigma_X - \sigma_Y}{2}\big)^2$. $T_{\max}$ intervient dans les critères de rupture (ou de fluage) d'un matériau.

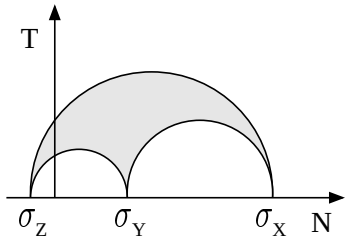


FIGURE 8

▷ Lois linéaires entre déformations et contraintes

Dans les matrices $[\boldsymbol{\epsilon}]$, ou $[\dot{\boldsymbol{\epsilon}}]$, et $[\boldsymbol{\sigma}]$ apparaissent une composante "isotrope" (scalaire) et une composante "quadrupolaire". La proportionnalité entre déformations et contraintes respecte ce découpage en vertu de l'invariance des lois physiques.

Pour les effets d'**élasticité** elle s'écrit $[\boldsymbol{\sigma}]_{\text{isotrope}} = \chi^{-1}\,\mathrm{tr}[\boldsymbol{\epsilon}]\,\mathbf{I}$ et $[\widetilde{\boldsymbol{\sigma}}] = 2\mu\,[\widetilde{\boldsymbol{\epsilon}}]$, soit :

$$\sigma_{ij} = 2\mu\,\epsilon_{ij} + \big(\chi^{-1} - \tfrac{2\mu}{3}\big)\,\textstyle\sum_k \epsilon_{kk}\,\delta_{ij} \quad ; \quad \text{par exemple} \quad \sigma_{xx} = \big(\chi^{-1} + \tfrac{4\mu}{3}\big)\,\epsilon_{xx} - \tfrac{2\mu}{3}\,(\epsilon_{yy} + \epsilon_{zz}) \ .$$

La définition de la **compressibilité** χ est bien connue (*cf.* $\chi = -\frac{1}{V}\,\frac{\delta V}{\delta P}$). Celle du **module de rigidité** μ est illustrée sur la figure 9a pour un cisaillement simple ($\overrightarrow{\psi} = \epsilon y \hat{x}$ d'où $[\widetilde{\boldsymbol{\epsilon}}] = \frac{1}{2}\begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix}$) ; la force

surfacique est alors $\frac{\overrightarrow{\mathrm{dF}}}{\mathrm{dS}} = \mu\epsilon\hat{x}$. On peut aussi exprimer les déformations en fonction des contraintes ; les relations $\epsilon_{xx} = \frac{1}{E}\,(\sigma_{xx} - \nu(\sigma_{yy} + \sigma_{zz}))$, $\sigma_{yy} = \cdots$ et $\sigma_{zz} = \cdots$ définissent le **module d'Young** E et le **coefficient de Poisson** ν. La figure 9b représente l'étirement d'une barre de section S par une force F : $\frac{\delta l}{l} = \epsilon_{xx} = \frac{1}{E}\,\frac{F}{S}$ si $\sigma_{yy} = \sigma_{zz} = 0$. La variation relative de volume étant alors $\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} = \frac{1}{E}\,(1-2\nu)\,\frac{F}{S}$, on en déduit $\chi = \frac{3(1-2\nu)}{E}$, d'où $\nu < \frac{1}{2}$. Enfin $\mu = 0$ caractérise un **fluide**. Pour les effets

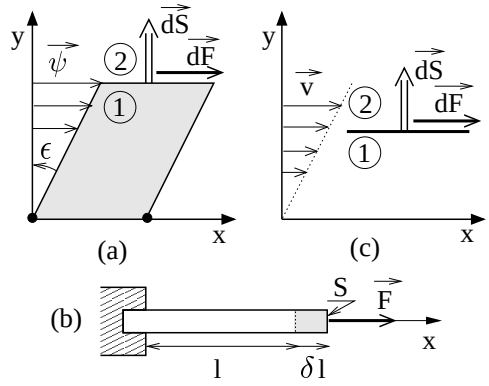


FIGURE 9

de viscosité les contraintes sont proportionnelles aux vitesses de déformation, ce qui se traduit par le remplacement de ψ_i par v_i dans les lois de l'élasticité : $[\boldsymbol{\sigma}]_{\text{isotrope}} = \zeta\,\mathrm{div}\vec{v}\,\mathbf{I}$ et $[\widetilde{\boldsymbol{\sigma}}] = 2\eta\,[\,\widetilde{\dot{\boldsymbol{\epsilon}}}\,]$, soit :

$$\sigma_{ij} = \eta\,(\partial_i v_j + \partial_j v_i) + \left(\zeta - \tfrac{2\eta}{3}\right)\,\mathrm{div}\vec{v}\,\delta_{ij}\ .$$

La signification du **coefficient de viscosité** η (seul coefficient non nul si le fluide est incompressible) est illustrée sur la figure 9c : $\frac{\overrightarrow{\mathrm{dF}}}{\mathrm{dS}} = \eta\,(\partial_y v_x)\,\hat{x}$ est proportionnel au gradient de vitesse. La relation pour la pression isotrope totale $\delta p = -\left(\chi^{-1}\,\frac{\delta V}{V} + \zeta\,\frac{\mathrm{d}}{\mathrm{dt}}\left(\frac{\delta V}{V}\right)\right)$ (si $\eta = 0$) montre que la présence du coefficient de **seconde viscosité** $\zeta > 0$ traduit, hors de l'équilibre, une avance temporelle de δp sur δV ; cet effet concerne les gaz non monoatomiques (*cf.* cours de physique).

4.3.2 Optique matricielle des systèmes centrés

On considère des systèmes présentant une symétrie de révolution autour de l'axe optique Oz et des rayons situés dans un plan méridien (plan contenant Oz), peu inclinés par rapport à Oz et intersectant les dioptres près de cet axe (**approximation de Gauss**). Un système centré est alors caractérisé par une correspondance linéaire entre rayons "entrant" et "sortant", qui justifie les constructions géométriques habituelles.

▷ Coordonnées d'un rayon ; formules de conjugaison (figure 10)

Dans un plan z_1 situé dans un milieu d'indice n_1, un rayon méridien est défini par une coordonnée spatiale transverse x_1 et une coordonnée "angulaire" $p_1 = n_1\alpha_1$, où α_1 est sa pente (algébrique), donc par un vecteur $\underline{R} = \begin{pmatrix} x_1 \\ p_1 \end{pmatrix}$. Les vecteurs tels que $x_1 = \alpha_1 q_1$, avec $q_1 = z_1 - z_{\mathrm{S}_1}$ fixé correspondent aux rayons issus de (ou allant vers) une source S_1 réelle ou virtuelle sur l'axe et d'abscisse z_{S_1} ; si la source est hors de l'axe $x_1 = x_{\mathrm{S}_1} + \alpha_1 q_1$. Quand un rayon passe d'un plan z_1 à un plan z_2, dans un milieu d'indice n_2, après avoir traversé des dioptres, le changement de ses coordonnées est donné par une **matrice de transfert** $\mathbf{M}(z_2, z_1)$ de déterminant 1 (démonstration ci-dessous) :

$$\begin{pmatrix} x_2 \\ n_2\alpha_2 \end{pmatrix} = \mathbf{M}(z_2, z_1) \begin{pmatrix} x_1 \\ n_1\alpha_1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ n_1\alpha_1 \end{pmatrix} \quad \text{avec } ad - bc = 1 \ .$$

On en déduit immédiatement que les rayons $x_1 = \alpha_1 q_1$ issus de S_1 sont transformés en rayons $x_2 = \alpha_2 q_2$ issus de S_2 avec :

$$\frac{q_2}{n_2} = \frac{a\frac{q_1}{n_1} + b}{c\frac{q_1}{n_1} + d} \quad \text{(relation de } \textbf{conjugaison} \text{ entre } S_1 \text{ et } S_2) \ ; \quad \frac{n_2\alpha_2}{n_1\alpha_1} = c\frac{q_1}{n_1} + d \ .$$

Par exemple (*cf.* matrices particulières ci-dessous) : $n_2 q_2^{-1} - n_1 q_1^{-1} = c$ si $a = d = 1$ et $b = 0$; $q_1 q_2 = n_1 n_2 c^{-2}$ si $a = d = 0$. On verra que $\dfrac{n_1\alpha_1}{n_2\alpha_2}$ est le grandissement γ_{S_2/S_1} (relation de Lagrange Helmholtz).

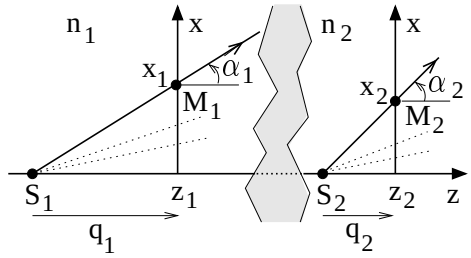


FIGURE 10

DÉMONSTRATION (figure 11a,b). La propagation rectiligne dans un milieu d'indice n constant et les lois de Descartes entraînent, pour une simple translation de Δz et pour un dioptre :

$$\mathbf{M}_{\text{ translation}} = \begin{pmatrix} 1 & \frac{\Delta z}{n} \\ 0 & 1 \end{pmatrix} \quad , \quad \mathbf{M}_{\text{ dioptre}} = \begin{pmatrix} 1 & 0 \\ -\mathcal{C} & 1 \end{pmatrix} \quad (\mathcal{C} = \frac{n' - n}{R}) \ .$$

(Pour le dioptre $x' = x$ et $n\left(\alpha + \dfrac{x}{R}\right) = n'\left(\alpha' + \dfrac{x}{R}\right)$ où $R = \overline{SC}$). Le produit de telles matrices est une matrice de déterminant 1.

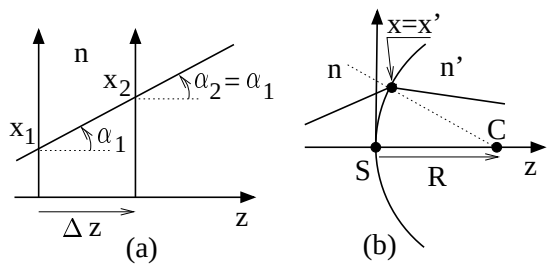


FIGURE 11

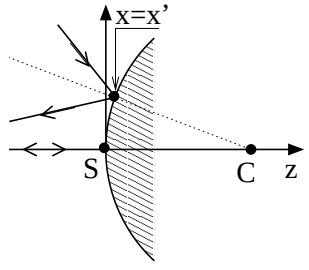


FIGURE 12

REMARQUES. 1) Si les rayons se dirigent vers les $z < 0$, les résultats ci-dessus restent valables à condition de remplacer $n\alpha$ par $(-n)\,\alpha$. Par exemple, pour le miroir de la figure 12, $\alpha + \frac{x}{R} = -\left(\alpha' + \frac{x}{R}\right)$ et $\mathcal{C} = -\frac{2}{R}$. 2) Si les rayons ne sont pas dans des plans méridiens, les formules ci-dessus s'appliquent aux projections des rayons sur de tels plans. 3) $\det \mathbf{M} = 1$ entraine $n_1\,\mathrm{d}x_1\,\mathrm{d}\alpha_1 = n_2\,\mathrm{d}x_2\,\mathrm{d}\alpha_2$ et donc la conservation de l'**étendue optique** (*cf.* section 3.2.5).

▷ Matrices de transfert particulières

Soient n et n' les indices des milieux d'entrée et de sortie d'un système optique, et O_e et O_s des origines dans chacun de ces milieux. La figure 13 représente un **système afocal** (avec $n = n' = 1$) ; la forme générale de la matrice est alors :

$$\mathbf{M}_{O_sO_e} = \begin{pmatrix} \gamma & b \\ 0 & \gamma^{-1} \end{pmatrix} \quad (\gamma \text{ indépendant de } O_e \text{ et } O_s) \ .$$

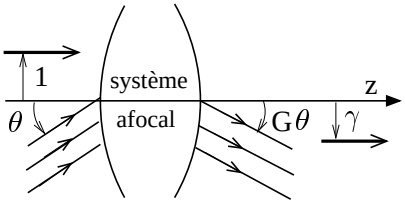


FIGURE 13

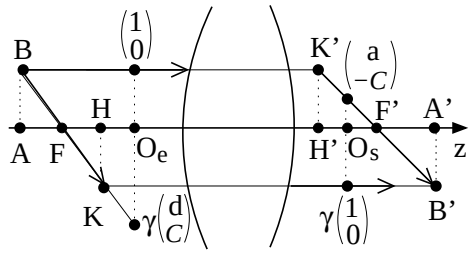


FIGURE 14

La relation angulaire $\alpha' = \gamma^{-1}\alpha$ donne le **grossissement** $G = \gamma^{-1}$. La figure 14 représente, pour un **système focal**, le rayon émergent $\begin{pmatrix} a \\ -\mathcal{C} \end{pmatrix}$ associé à l'incident $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et l'incident $\gamma \begin{pmatrix} d \\ \mathcal{C} \end{pmatrix}$ associé à l'émergent $\gamma \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, la forme générale de $\mathbf{M}$ et $\mathbf{M}^{-1}$ étant :

$$\mathbf{M}_{O_sO_e} = \begin{pmatrix} a & b \\ -\mathcal{C} & d \end{pmatrix} \quad (\mathcal{C} \textbf{ convergence}) \quad \mathbf{M}^{-1}_{O_sO_e} = \begin{pmatrix} d & -b \\ \mathcal{C} & a \end{pmatrix} \ ;$$

a dépend du choix de O_s (par exemple $a = 1$ pour $O_s \equiv H'$ point principal image et $a = 0$ pour $O_s \equiv F'$ foyer image) et d dépend de celui de O_e (par exemple $d = 1$ pour $O_e \equiv H$ point principal objet et $d = 0$ pour $O_e \equiv F$ foyer objet) ; par contre la convergence $\mathcal{C}$, qui ne dépend ni de O_e ni de O_s, est une caractéristique du système. On voit aussi que $\mathcal{C} = \dfrac{n'}{\overline{H'F'}} = -\dfrac{n}{\overline{HF}}$ et on déduit de $\det \mathbf{M} = 1$ que $\mathbf{M}_{H'H} = \begin{pmatrix} 1 & 0 \\ -\mathcal{C} & 1 \end{pmatrix}$ et $\mathbf{M}_{F'F} = \begin{pmatrix} 0 & \mathcal{C}^{-1} \\ -\mathcal{C} & 0 \end{pmatrix}$. Enfin la géométrie des triangles semblables FAB et FHK d'une part et F'H'K' et F'A'B' d'autre part donne les **relations de Newton** $\gamma_{\mathrm{A'/A}} = \dfrac{\overline{A'B'}}{\overline{AB}} = \dfrac{\overline{FH}}{\overline{FA}} = \dfrac{\overline{F'A'}}{\overline{F'H'}}$ ($\gamma_{\mathrm{A'/A}} = \gamma = -1$ sur la figure).

Condition de conjugaison. La figure 15 illustre la marche de rayons particuliers pour tout couple $O_e \equiv A$ et $O_s \equiv A'$ tel que $b = 0$; elle montre que $b = 0$, qui entraîne $x' = ax$ quel que soit α, est la condition de conjugaison de A et A' et que $a = \gamma_{\mathrm{A'/A}}$ est le grandissement. Pour des rayons passant par A, $\det \mathbf{M} = ad = 1$ conduit à la **relation de Lagrange Helmholtz** : $n'\theta' = \gamma^{-1}_{\mathrm{A'/A}} n\theta$.

Association de deux systèmes focaux de foyers F_i et F'_i. On suppose $n = 1$ en dehors des deux systèmes. Le calcul du produit $\mathbf{M}_{F'_2F_2}\mathbf{M}_{F_2F'_1}\mathbf{M}_{F'_1F_1}$ donne $\mathcal{C} = \mathcal{C}_1\mathcal{C}_2\,\overline{F'_1F_2}$; donc $\mathcal{C} = 0$ si $F'_1 \equiv F_2$, et alors $G = -\frac{\mathcal{C}_2}{\mathcal{C}_1}$.

Cavité laser (figure 16) : $R_1 = \overline{S_1C_1}$, $R_2 = -\overline{S_2C_2}$ et $l = \overline{S_1S_2}$. Pour un rayon incident sur le miroir 2, après deux réflexions la matrice de transfert s'écrit :

$$\mathbf{M}_{S_2S_2} = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{2}{R_1} & 1 \end{pmatrix} \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{2}{R_2} & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

La cavité est dite stable si les rayons restent proches de l'axe donc si les éléments de $\mathbf{M}^n$ restent bornés ; les valeurs propres de $\mathbf{M}$ doivent donc satisfaire $|\lambda| \leq 1$. Comme $\det \mathbf{M} = \lambda_1\lambda_2 = 1$, ceci n'est possible que si $\lambda = e^{\pm i\varphi}$, soit $|\mathrm{tr}\mathbf{M}| = |a+d| = |2\cos\varphi| < 2$. Remarque : si $\lambda = \pm 1$ est une valeur propre double et si $\mathbf{M}$ n'est pas diagonalisable il y a **instabilité** : $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & nb \\ 0 & 1 \end{pmatrix}$ (cas par exemple de deux miroirs plans).

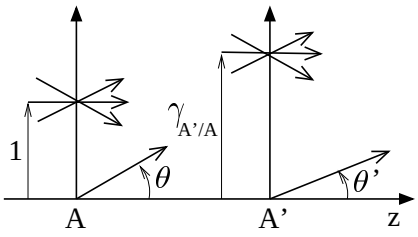


FIGURE 15

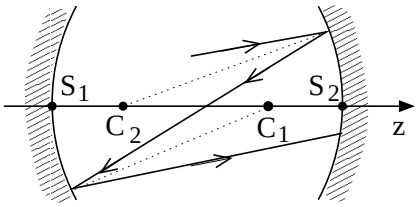


FIGURE 16

▷ Optique matricielle et principe de Fermat (figure 10)

Le chemin optique $\mathcal{L}(x_1, x_2)$ joignant deux points M_1 et M_2 des plans z_1 et z_2 vérifie :

$$\mathcal{L}(x_1, x_2) - \mathcal{L}(0,0) = \frac{1}{2b}\left(ax_1^2 + dx_2^2 - 2x_1x_2\right) .$$

DÉMONSTRATION. Que $\mathcal{L}(x_1, x_2) - \mathcal{L}(0,0)$ soit une forme quadratique, à l'ordre le plus bas, résulte simplement de la symétrie de révolution du système autour de l'axe z. Pour montrer que les coefficients a, b et d sont bien les éléments de la matrice $\mathbf{M}$ introduite précédemment, on calcule $n_1\alpha_1$ et $n_2\alpha_2$ en fonction de x_1 et x_2 à l'aide de la formule $\mathrm{d}\mathcal{L} = b^{-1}\left(\mathrm{d}x_2 - x_1\right)\mathrm{d}x_2 + b^{-1}\left(ax_1 - x_2\right)\mathrm{d}x_1 = n_2\alpha_2\,\mathrm{d}x_2 - n_1\alpha_1\,\mathrm{d}x_1$, cas particulier de $\mathrm{d}L(A,B) = n_B\,\hat{u}_B \cdot \overrightarrow{\mathrm{d}r}_B - n_A\,\hat{u}_A \cdot \overrightarrow{\mathrm{d}r}_A$ (*cf.* section 3.2.3). Ce calcul montre que la propriété $\det \mathbf{M} = 1$ est liée à l'existence d'un principe d'extremum.

Cas particuliers : $\mathcal{L}(x_1, x_2) - \mathcal{L}(0,0)$ vaut $\dfrac{1}{2}\dfrac{n}{\Delta z}\left(x_2 - x_1\right)^2$ pour $\mathbf{M}_{\text{translation}}$ et $-\mathcal{C}xx'$ pour $\mathbf{M}_{F'F}$. Pour $\mathbf{M}_{H'H}$ ou $\mathbf{M}$ $_{\text{dioptre}}$ et des points conjugués avec $x = x'$ (sinon il n'y a pas de rayon joignant ces points), on obtient "$\dfrac{0}{0}$" ; pour lever l'indétermination on note que $ad - bc = 1$, donc $a = 1 + bc$ si $d = 1$, et on obtient $-\dfrac{1}{2}\mathcal{C}x^2$.

4.3.3 Relativité d'Einstein et quadrivecteurs

Rappelons que le groupe de symétrie (**groupe de Lorentz**), associé aux changements de référentiels inertiels, correspond aux transformations linéaires qui laissent invariante la métrique $c^2T^2 - \overrightarrow{R}^2$, où T et $(X, Y, Z) \equiv \overrightarrow{R}$ sont les composantes d'un intervalle

d'espace-temps. Une **transformation de Lorentz pure** de **rapidité** φ ($\tanh\varphi = \dfrac{V}{c}$, $\cosh\varphi = \left(1 - \dfrac{V^2}{c^2}\right)^{-\frac{1}{2}}$) selon l'axe x s'écrit :

$$cT' = cT\,\cosh\varphi - X\,\sinh\varphi\ ,\ X' = X\,\cosh\varphi - cT\,\sinh\varphi\ ,\ Y' = Y \text{ et } Z' = Z\ .$$

Si la transformation se fait dans la direction $\hat{n}$, il faut remplacer X par $\overrightarrow{R}\cdot\hat{n}$ et le couple (Y, Z) par $\overrightarrow{R}_\perp = \overrightarrow{R} - (\overrightarrow{R}\cdot\hat{n})\,\hat{n} = \overrightarrow{R} - \overrightarrow{R}_\parallel$; en particulier pour une **transformation infinitésimale** :

$$cT' = cT - \vec{\epsilon}\cdot\overrightarrow{R} \quad , \quad \overrightarrow{R}' = \overrightarrow{R} - cT\,\vec{\epsilon} \quad \text{avec } \vec{\epsilon} = \frac{V}{c}\,\hat{n}\ .$$

L'invariance des lois physiques par ces transformations infinitésimales entraîne leur invariance par les "transformations finies" (en vertu de la loi de groupe).

▷ Classification des évènements et des vitesses

La figure 17 représente les courbes $c^2T^2 - \overrightarrow{R}^2 = \mathrm{C}^{\mathrm{ste}}$ dans le demi-plan $(cT, |\overrightarrow{R}|)$. Chaque courbe montre les valeurs possibles de T et $|\overrightarrow{R}|$, pour un intervalle d'espace-temps donné, quand on change de référentiel.

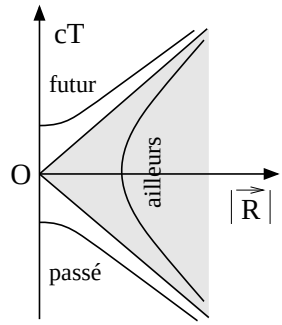


FIGURE 17

On voit sur cette figure que la notion de simultanéité, $T = 0$, de deux évènements distincts n'est pas invariante. Quant à la notion de **causalité**, qui pour être valable demande que le signe de T soit invariant, on voit qu'elle n'est applicable qu'à des évènements tels que $c^2T^2 - \overrightarrow{R}^2 > 0$ ou $c^2T^2 - \overrightarrow{R}^2 = 0$; dans ce dernier cas T ne peut changer de signe car si $T = 0$ dans un référentiel alors $\overrightarrow{R} = 0$ également et les deux évènements sont confondus. Pour un évènement origine O on distingue ainsi les évènements qui sont dans son **futur** ou son **passé**, et qui peuvent être reliés à lui par causalité, et ceux tels que $c^2T^2 - \overrightarrow{R}^2 < 0$, qui dans tout référentiel se produisent toujours "ailleurs" ($|\overrightarrow{R}| \neq 0$) et antérieurement ou postérieurement suivant le référentiel. Dans l'espace à quatre dimensions (cT, X, Y, Z) les futur et passé de O sont à l'intérieur d'un cône appelé **cône de lumière**.

Si on pose $\vec{v} = T^{-1}\,\overrightarrow{R}$ ($\vec{v}$ est la vitesse moyenne d'un point qui, par la pensée, "joindrait" les deux évènements), la relation $c^2T^2 - \overrightarrow{R}^2 = c^2T^2\left(1 - \dfrac{\vec{v}^2}{c^2}\right)$ montre que les vitesses se classent de manière invariante en trois catégories : $|\vec{v}| < c$, $|\vec{v}| = c$ (vitesse c "limite" des précédentes) et $|\vec{v}| > c$. Pour la première seulement il existe un référentiel inertiel $\mathcal{R}_0$ dans lequel les évènements se produisent au même endroit ($\overrightarrow{R}_0 = 0$).

REMARQUES. 1) La **loi de transformation des vitesses**, obtenue en divisant X', Y' et Z' par T', s'écrit vectoriellement

$$\vec{v}' = \left(1 - \tfrac{\vec{v}\cdot\overrightarrow{V}}{c^2}\right)^{-1}\left(\vec{v}_{\parallel} - \overrightarrow{V} + \vec{v}_{\perp}\sqrt{1 - \tfrac{V^2}{c^2}}\right) \quad \text{d'où } v' = \tfrac{v-V}{1-\frac{vV}{c^2}} \text{ à une dimension .}$$

Si on remplace $\vec{v}$ et $\overrightarrow{V}$ par les vitesses $\vec{v}_2$ et $\vec{v}_1$ de deux mobiles, le vecteur $\vec{v}'$ représente la **vitesse relative** $\vec{v}_{21}$ de 2 par rapport à 1 et $v_{21}^2 = \left(1 - \tfrac{(\vec{v}_1\cdot\vec{v}_2)^2}{c^2}\right)^{-2}\left((\vec{v}_2 - \vec{v}_1)^2 - \tfrac{1}{c^2}\,(\vec{v}_1 \wedge \vec{v}_2)^2\right)$ est une quantité invariante. On a $\vec{v}_{21} = \vec{v}_2 - \vec{v}_1$ dans l'approximation galiléenne. 2) Dans l'espace des vitesses (v_x, v_y, v_z), intérieur de la sphère $|\vec{v}| < c$, les transformations ci-dessus, et les rotations, transforment les droites en droites et conservent la "distance" Φ_{21} de deux "points" définie comme la rapidité relative ($v_{21} = c\,\tanh\Phi_{21}$). Ceci définit une géométrie équivalente à la **géométrie non euclidienne de Lobatchevski**. On vérifie que les points $|\vec{v}| = c$ sont les points "à l'infini" ; deux droites étant par définition parallèles si elles se coupent à l'infini, la figure 18 montre que par tout point M passe deux droites MA et MB parallèles à une droite donnée AB.

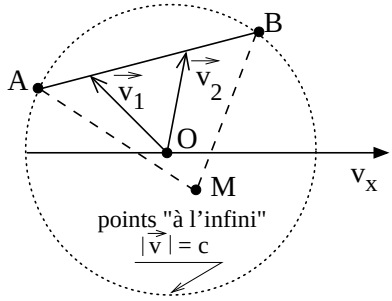


FIGURE 18

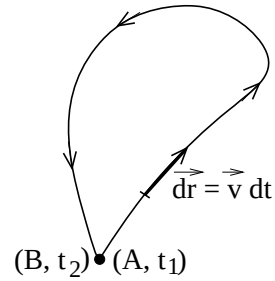


FIGURE 19

▷ Temps propre ; ralentissement des horloges ; effet Doppler

Si dans un référentiel $\mathcal{R}$, un point matériel se déplace de $\overrightarrow{\mathrm{d}r}$ dans le temps $\mathrm{d}t$ (à la vitesse $\vec{v}$), l'invariance de $c^2\,\mathrm{d}t^2 - |\overrightarrow{\mathrm{d}r}|^2$ montre que le temps $\mathrm{d}\tau$ écoulé dans le référentiel inertiel $\mathcal{R}_0$ où $\overrightarrow{\mathrm{d}r}_0 = 0$ (référentiel qui change avec t), est tel que $\mathrm{d}\tau^2 = \mathrm{d}t^2\left(1 - \dfrac{v^2}{c^2}\right)$. On appelle temps propre entre deux "positions" (A, t_1) et (B, t_2) du mobile dans $\mathcal{R}$ l'invariant :

$$\tau = \int_{(A,t_1)}^{(B,t_2)} \sqrt{1 - \frac{v^2}{c^2}}\,\mathrm{d}t \qquad \left(\mathrm{d}\tau = \mathrm{d}t\sqrt{1 - \frac{v^2}{c^2}}\right) .$$

Si (A, t_1) et (B, t_2) sont fixés τ est maximum pour un **mouvement inertiel**. Cette propriété (exploitée à la section 9.2.1) est évidente si on se place dans le référentiel inertiel où A et B coïncident (figure 19) ; en effet le mouvement inertiel de A à $B \equiv A$ correspond alors à $v = 0$ et $\tau = t_2 - t_1$ est maximum.

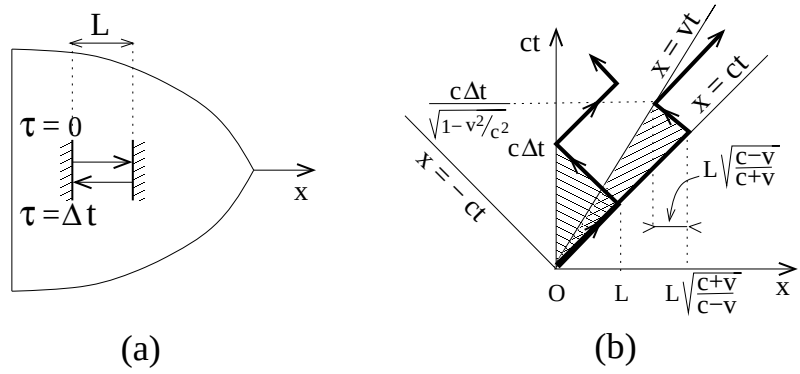


FIGURE 20

REMARQUES. 1) La figure 20b représente dans l'espace-temps (ct, x) d'un référentiel l'aller-retour d'un signal lumineux entre deux miroirs, distants de L dans leur référentiel propre (fusée de la figure 20a), lorsque la fusée est soit au repos soit a la vitesse v dans ce référentiel. Les aires hachurées sont égales car une transformation de Lorentz (ici active) préserve $c^2t^2 - x^2 = (ct + x)(ct - x)$. Le système des miroirs est un exemple d'horloge (avec période propre $\Delta t = 2L/c$) qui, dans le référentiel (ct, x), manifeste un **ralentissement** (période $\Delta t \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$) quand elle est en mouvement. 2) Même si le **temps** (propre) **de vie** d'une particule est petit, $\tau \simeq 10^{-10}$ s par exemple, celle ci peut parcourir dans $\mathcal{R}$ des distances $v\tau\left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$ arbitrairement grandes quand v s'approche de c. 3) L'accélération du mobile dans $\mathcal{R}$ est $\vec{a} = \frac{\mathrm{d}\vec{v}}{\mathrm{d}t}$; en posant $\overrightarrow{V} = \vec{v}(t)$ et $\vec{v} = \vec{v}(t + \mathrm{d}t)$, on déduit de la loi de transformation des vitesses que son **accélération propre** (dans $\mathcal{R}_0$) $\frac{\vec{v}\,'}{\mathrm{d}\tau}$ est $\vec{a}^{(0)} = \left(1 - \frac{V^2}{c^2}\right)^{-\frac{3}{2}} \left(\vec{a}_{\|} + \vec{a}_{\perp}\sqrt{1 - \frac{V^2}{c^2}}\right)$. 4) **Effet Doppler**. L'étude a été faite à la section 3.2.3. Si la vitesse de propagation du signal est $c\hat{u}$ (lumière) et celle du détecteur $v_D = 0$, la relation entre fréquences s'écrit $\nu_{\mathrm{D}} = \nu_{\mathrm{S}}\left(1 - \frac{\vec{v}_S \cdot \hat{u}}{c}\right)^{-1}\left(1 - \frac{v_S^2}{c^2}\right)^{\frac{1}{2}}$; si $\vec{v}_S$ est parallèle à $\hat{u}$ cette relation devient $\nu_D = \nu_S\, e^{\varphi}$ où φ est la rapidité de la source (*cf.* section 1.1.2.).

▷ Quadrivecteurs

On appelle ainsi tout ensemble $\underline{A}$ de quatre composantes $(A_0, \overrightarrow{A} \equiv \{A_x, A_y, A_z\})$ qui se transforme comme $\underline{X} = (cT, \overrightarrow{R})$. Le carré de la norme d'un quadrivecteur, défini par $A_0^2 - \overrightarrow{A}^2$, est un invariant (scalaire de Lorentz) de même que le produit scalaire $\underline{A} \cdot \underline{B} \overset{\text{def}}{=} A_0B_0 - \overrightarrow{A} \cdot \overrightarrow{B}$ (conséquence de l'invariance de $(A_0 + B_0)^2 - (\overrightarrow{A} + \overrightarrow{B})^2$). Un quadrivecteur est du genre temps, lumière ou espace selon que ce carré est positif, nul ou négatif ; dans le premier cas le signe de A_0 est un invariant. Un exemple est la **quadrivitesse** (de norme 1)

$$\underline{U} = \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\overrightarrow{\mathrm{d}r}}{c\,\mathrm{d}\tau}\right) = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \left(1, \frac{\vec{v}}{c}\right)$$

obtenue en divisant $\mathrm{d}\underline{X} = (c\,\mathrm{d}t, \overrightarrow{\mathrm{d}r})$ par le scalaire $c\,\mathrm{d}\tau$. On construit à partir de $\underline{U}$ le **quadrivecteur énergie quantité de mouvement** $\underline{P}$ d'une particule ponctuelle par multiplication par le scalaire mc (la masse m est supposée invariante) :

$$\underline{P} = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} (mc, m\overrightarrow{v}) = \left(\frac{E}{c}, \overrightarrow{p}\right) \quad \text{d'où} \quad \overrightarrow{v} = \frac{c^2\overrightarrow{p}}{E}\ .$$

AUTRES EXEMPLES. 1) On construit de même le **quadrivecteur courant** $\underline{J}$ d'une distribution de charges, de densité ρ_0 dans son référentiel propre et animée d'une vitesse $\vec{v}$ dans $\mathcal{R}$:

$$\underline{J} = \rho_0 \, \underline{U} = \left(\rho, \tfrac{\vec{j}}{c}\right) \quad \text{avec } \vec{j} = \rho\vec{v} \text{ et } \rho = \rho_0 \left(1 - \tfrac{v^2}{c^2}\right)^{-\frac{1}{2}} ;$$

ρ est la densité volumique de charges dans $\mathcal{R}$ car (*cf.* section 3.3.1, figure 3) la charge $\rho_0 \, \mathrm{d}^3 r_0$ (supposée invariante) se retrouve à un instant donné t dans $\mathcal{R}$ dans le volume $\mathrm{d}^3 r = \mathrm{d}^3 r_0 \sqrt{1 - \frac{v^2}{c^2}}$. 2) Si pour une onde plane $\exp -i(\omega t - \vec{k}\cdot\vec{r})$ on suppose que le déphasage $\omega T - \vec{k}\cdot\overrightarrow{R}$ est un invariant, on en déduit que

$$\underline{k} = \left(\tfrac{\omega}{c}, \vec{k}\right) \quad (\textbf{“quadrifréquence”})$$

est un quadrivecteur. (Démonstration : $\omega T - \vec{k}\cdot\overrightarrow{R} = \omega(T' + \frac{1}{c}\,\vec{\epsilon}\cdot\overrightarrow{R}') - \vec{k}\,(\overrightarrow{R}' + cT'\vec{\epsilon})$ s'écrit aussi $\omega' T' - \vec{k}'\cdot\overrightarrow{R}'$ avec $\frac{\omega'}{c} = \frac{\omega}{c} - \vec{\epsilon}\cdot\vec{k}$ et $\vec{k}' = \vec{k} - \frac{\omega}{c}\,\vec{\epsilon}$.) Les **relations de Planck-Einstein et de de Broglie** $E = \hbar\omega$ et $\vec{p} = \hbar\vec{k}$, équivalentes à $\underline{p} = \hbar\underline{k}$, sont donc invariantes. 3) L'opérateur **gradient**

$$\underline{\partial} = (\tfrac{1}{c}\,\partial_t, -\overrightarrow{\nabla})$$

est aussi un quadrivecteur ; appliqué à une onde plane, $i\underline{\partial}$ est la multiplication par $\underline{k}$.

▷ Mécanique des points ; défaut de masse ; collisions ; forces

Elle repose sur la définition de $\underline{P}$ qui entraîne $E^2 - c^2\,\vec{p}^{\,2} = m^2c^4$; pour les particules de **masse nulle** $E = c\,|\vec{p}|$. Si $m \neq 0$, mc^2 est l'énergie au repos et $E - mc^2$ ($\simeq \frac{1}{2}\,mv^2$ si $v \ll c$) est l'**énergie cinétique**.

EXEMPLES. En chimie ou physique nucléaire, on pose $mc^2 = m^*c^2 + E^{\text{ liaison}}$, où m^* est la somme des masses des constituants isolés et $E^{\text{ liaison}}$ l'**énergie de liaison** ; par exemple pour l'atome d'hydrogène : $m_H c^2 = (m_e + m_p)c^2 - 13{,}6\,\text{ev}$. La variation d'énergie cinétique lors d'une réaction (“énergie libérée”) est alors :

$$\begin{aligned}
\Delta E^{\text{ cin}} &= \textstyle\sum_f (E_f - m_f c^2) - \sum_i (E_i - m_i c^2) \\
&= -c^2 \left(\textstyle\sum_f m_f - \sum_i m_i\right) \text{ (conservation de l'énergie)} \\
&= -\left(\textstyle\sum_f E_f^{\text{liaison}} - \sum_i E_i^{\text{liaison}}\right) \text{ si les constituants restent inchangés .}
\end{aligned}$$

Elle est positive si $\sum_f m_f < \sum_i m_i$ (**défaut de masse**, c.a.d. en chimie états finaux davantage liés).

Collisions entre particules. Elles obéissent à la loi de conservation invariante (égalité entre quadrivecteurs) :

$$\Sigma_f \underline{P}_f = \Sigma_i \underline{P}_i = \underline{P} \ .$$

Comme $E_i \geq c\,|\overrightarrow{p}_i|$ on a $\underline{P}^2 > 0$ sauf si les particules incidentes sont de masses nulles et ont des vitesses parallèles. Alors il existe un référentiel $\mathcal{R}^{cm}$ où la somme $\vec{p}^{\,cm}$ des quantités de mouvement est nulle (sa vitesse dans $\mathcal{R}$ est $\vec{v}^{\,cm} = \frac{\Sigma_i c^2 \vec{p}_i}{\Sigma_i E_i}$). On en déduit :

$$\underline{P}^{\,cm} = (\frac{E^{cm}}{c}, \vec{0}) \quad ; \quad \frac{(E^{cm})^2}{c^2} = \underline{P}^2 = (\sum_i \underline{P}_i)^2 \ .$$

EXEMPLES. 1) Pour la collision d'une particule (2) sur une particule cible (1) au repos dans $\mathcal{R}$:

$$(E^{cm})^2 = (m_1c^2 + E_2)^2 - \vec{p}_2^{\,2}\,c^2 = (m_1^2 + m_2^2)\,c^4 + 2m_1c^2E_2\ .$$

La quantité $E^{cm} - (m_1 + m_2)\,c^2$ est l'énergie disponible pour la création de particules ; la figure 21a représente ce qu'on observe dans $\mathcal{R}$ au **seuil de production**. A basse énergie on retrouve (petit calcul) l'expression galiléenne de l'énergie disponible $\frac{1}{2}\,\frac{m_1m_2}{m_1+m_2}\,\vec{v}_2\,^2$ (proportionnelle à $E_2^{\rm cin}$; *cf.* section 3.3.2) ; à très haute énergie elle vaut $\sqrt{2m_1c^2E_2}$ (proportionnalité à $\sqrt{E_2^{\rm cin}}$ qui montre l'intérêt d'utiliser plutôt des anneaux de collision).

2) Dans une collision élastique, l'angle de diffusion de la particule incidente (2) s'obtient en éliminant les paramètres de la particule cible (1) après le choc : $\underline{P}_1'^2 = m_1^2c^2 = (\underline{P}_1 + \underline{P}_2 - \underline{P}_2')^2$; par exemple si (2) est un photon ($m_2 = 0$) qui diffuse sous un angle θ, la relation s'écrit

$$m_1^2c^4 = (m_1c^2 + E_2 - E_2')^2 - (\vec{p}_2 - \vec{p}_2\,')^2\,c^2 \quad \text{ou} \quad 2m_1c^2\,(E_2 - E_2') = 2E_2E_2'\,(1 - \cos\theta)\ ,$$

ou encore :

$$\lambda' - \lambda = \tfrac{h}{m_1c}\,(1 - \cos\theta) \quad (\textbf{effet Compton}\,;\ \text{figure 21b}).$$

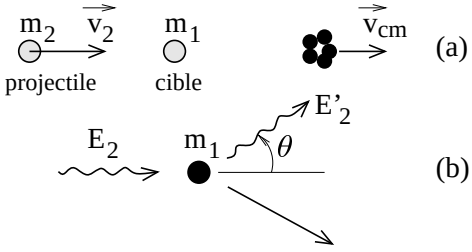


FIGURE 21

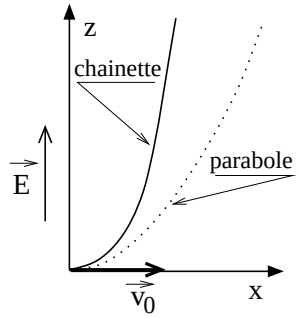


FIGURE 22

Force et quadriforce. La différentiation de $E^2 - \vec{p}^{\,2}c^2 = m^2c^4$ montre que $\mathrm{d}E = \vec{v}\cdot\overrightarrow{\mathrm{d}p}$; en gardant la définition usuelle de la force on a :

$$\frac{\overrightarrow{\mathrm{d}p}}{\mathrm{d}t} = \overrightarrow{F} \quad \text{et} \quad \frac{\mathrm{d}E}{\mathrm{d}t} = \overrightarrow{F}\cdot\vec{v}\ .$$

En paramétrant la trajectoire d'espace temps par le temps propre τ, et en notant par un point la dérivation $\dfrac{\mathrm{d}}{\mathrm{d}\tau}$, ceci s'écrit aussi :

$$\underline{\dot{P}} = m\underline{\ddot{X}} = \underline{F} \quad \text{avec} \quad \underline{F} = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}\ \left(\frac{\overrightarrow{F}\cdot\vec{v}}{c}\,,\ \overrightarrow{F}\right)\ \textbf{quadriforce}\ .$$

EXEMPLE : **mouvement d'une particule chargée dans un champ électrique constant** $\overrightarrow{E} = E_z\,\hat{z}$. $mc^2\ddot{t} = qE_z\dot{z}$, $m\ddot{x} = m\ddot{y} = 0$ et $m\ddot{z} = qE_z\dot{t}$. Sans perte de généralité (réflexion laissée au lecteur), on peut choisir $x(0) = y(0) = z(0) = t(0) = 0$, $\dot{y}(0) = \dot{z}(0) = 0$ et $\dot{x}(0) = v_0$, d'où $\dot{t}(0) = \left(1 - \frac{v_0^2}{c^2}\right)^{-\frac{1}{2}} = \gamma_0$. L'intégration de $\ddot{t}$ donne la L.C. de l'énergie totale (cinétique et potentielle) $mc^2\dot{t} - qE_zz = mc^2\gamma_0$; celle de $\ddot{x}$ et de $\ddot{y}$ conduit à : $x(\tau) = \gamma_0v_0\tau$ et $y(\tau) = 0$; enfin $m\ddot{z} = \frac{qE_z}{mc^2}\,(qE_zz + mc^2\gamma_0)$ et les C.I. donnent $z(\tau) = -\frac{mc^2\gamma_0}{qE_z}\,\left(1 - \cosh\frac{qE_z}{mc}\,\tau\right)$ et $t(\tau) = \frac{m\dot{z}}{qE_z} = \frac{\gamma_0mc}{qE_z}\,\sinh\frac{qE_z}{mc}\,\tau$. Pour τ petit $t = \gamma_0\tau$, $x = v_0t$, $z = \frac{1}{2}\,\frac{qE_z}{\gamma_0\,m}\,t^2$ et la trajectoire est une parabole. Pour τ grand c'est une chaînette (figure 22) ; p_x restant constant, v_x tend vers zéro tandis que v_z tend vers c.

▷ Invariance de l'électromagnétisme

La **loi de force de Lorentz** $\overrightarrow{F} = q\,(\overrightarrow{E} + \vec{v} \wedge \overrightarrow{B})$ est invariante par transformation de Lorentz et la loi de transformation infinitésimale de $\overrightarrow{E}$ et $\overrightarrow{B}$ est :

$$\overrightarrow{E}\,' = \overrightarrow{E} + \vec{\epsilon} \wedge c\overrightarrow{B} (= \overrightarrow{E} + \overrightarrow{V} \wedge \overrightarrow{B}) \quad \text{et} \quad \overrightarrow{B}\,' = \overrightarrow{B} - \vec{\epsilon} \wedge \frac{1}{c}\overrightarrow{E}$$

(*cf.* démonstration 1). On en déduit facilement que $\overrightarrow{E} \cdot \overrightarrow{B}$ et $\overrightarrow{E}^{\,2} - c^2 \overrightarrow{B}^{\,2}$ sont invariants, et que $\overrightarrow{E}$ et $c\overrightarrow{B}$ se transforment comme $\overrightarrow{B}$ et $-\frac{1}{c}\overrightarrow{E}$. Cette correspondance, appelée relation de dualité, entraîne que la loi $\overrightarrow{F} = q_m \left(\overrightarrow{B} - \frac{1}{c^2}\,\vec{v} \wedge \overrightarrow{E}\right)$ est elle aussi invariante ; c'est la force que subirait une charge magnétique q_m s'il existait des monopôles magnétiques. On déduit aussi des lois de transformation de $\overrightarrow{E}$ et $\overrightarrow{B}$ que les **équations de Maxwell** dans le vide (*cf.* section 7.5.1), $\operatorname{div}\overrightarrow{E} = \dfrac{\rho}{\epsilon_0}$, $\overrightarrow{\operatorname{rot}}\,\overrightarrow{B} = \mu_0\,(\vec{j} + \epsilon_0 \partial_t \overrightarrow{E})$, $\operatorname{div}\overrightarrow{B} = 0$ et $\overrightarrow{\operatorname{rot}}\,\overrightarrow{E} + \partial_t \overrightarrow{B} = 0$ sont invariantes (*cf.* démonstration 2).

DÉMONSTRATION 1. Il suffit d'écrire la loi de force sous la forme $\overrightarrow{\mathrm{d}p} = q\,(\overrightarrow{E}\,\mathrm{d}t + \overrightarrow{\mathrm{d}r} \wedge \overrightarrow{B})$; comme $(\frac{\mathrm{d}E}{c}, \overrightarrow{\mathrm{d}p})$ est un quadrivecteur, la transformation de $\overrightarrow{\mathrm{d}p}$ à l'ordre $\vec{\epsilon}$ est :

$$\begin{aligned}\overrightarrow{\mathrm{d}p}\,' &= \overrightarrow{\mathrm{d}p} - \vec{\epsilon}\,\frac{1}{c}\,\mathrm{d}E = q\,\big(\overrightarrow{E}(\mathrm{d}t' + \vec{\epsilon}\cdot\frac{1}{c}\overrightarrow{\mathrm{d}r}\,') + (\overrightarrow{\mathrm{d}r}\,' + \vec{\epsilon} c\,\mathrm{d}t') \wedge \overrightarrow{B}\big) - \vec{\epsilon}\,\frac{1}{c}\,q\,\overrightarrow{E}\cdot\overrightarrow{\mathrm{d}r}\,' \\ &= q\,(\overrightarrow{E}\,'\,\mathrm{d}t' + \overrightarrow{\mathrm{d}r}\,' \wedge \overrightarrow{B}\,')\ ;\end{aligned}$$

l'identification des termes en $\mathrm{d}t'$ donne la loi de transformation de $\overrightarrow{E}$ et celle des termes en $\overrightarrow{\mathrm{d}r}\,'$, et l'utilisation de la formule du double produit vectoriel, celle de $\overrightarrow{B}$. L'intégration des relations $\frac{\mathrm{d}\overrightarrow{E}(\varphi)}{\mathrm{d}\varphi} = c\,\hat{n} \wedge \overrightarrow{B}(\varphi)$ et $\frac{\mathrm{d}\overrightarrow{B}(\varphi)}{\mathrm{d}\varphi} = -\hat{n} \wedge \frac{1}{c}\,\overrightarrow{E}(\varphi)$ (φ rapidité) donne, pour des transformations finies :
$\overrightarrow{E'_{\parallel}} = \overrightarrow{E}_{\parallel}$, $\overrightarrow{B'_{\parallel}} = \overrightarrow{B}_{\parallel}$, $\overrightarrow{E'_{\perp}} = \cosh\varphi\,\overrightarrow{E}_{\perp} + \sinh\varphi\,\hat{n} \wedge c\overrightarrow{B}_{\perp}$ et $\overrightarrow{B'_{\perp}} = \cosh\varphi\,\overrightarrow{B}_{\perp} - \sinh\varphi\,\hat{n} \wedge \frac{1}{c}\,\overrightarrow{E}_{\perp}$.

DÉMONSTRATION 2. On vient de voir que les lois de transformation de $\overrightarrow{E}$ et $\overrightarrow{B}$ sont telles que les quantités $(\frac{1}{c}\,\overrightarrow{E}\cdot\mathrm{d}\vec{r}, \overrightarrow{E}\,\mathrm{d}t + \mathrm{d}\vec{r} \wedge \overrightarrow{B})$ et $(\overrightarrow{B}\cdot\mathrm{d}\vec{r}, c\,\overrightarrow{B}\,\mathrm{d}t - \frac{1}{c}\,\mathrm{d}\vec{r} \wedge \overrightarrow{E})$ sont des quadrivecteurs. Comme $(\frac{1}{c}\,\partial_t, -\overrightarrow{\nabla})$ et $(c\,\mathrm{d}t, \mathrm{d}\vec{r})$ se transforment de la même façon, on en déduit que $(-\frac{1}{c}\,\overrightarrow{\nabla}\cdot\overrightarrow{E}, \frac{1}{c^2}\,\partial_t\overrightarrow{E} - \overrightarrow{\nabla} \wedge \overrightarrow{B})$ et $(-\overrightarrow{\nabla}\cdot\overrightarrow{B}, \frac{1}{c}\,\partial_t\overrightarrow{B} + \overrightarrow{\nabla} \wedge \frac{1}{c}\,\overrightarrow{E})$ sont aussi des quadrivecteurs. Les équations de Maxwell, qui expriment l'égalité de ces quadrivecteurs avec les quadrivecteurs $-\frac{1}{\epsilon_0 c}\,(\rho, \frac{1}{c}\,\vec{j})$ et $(0, \vec{0})$ (proportionnels aux courants électrique et magnétique), sont donc invariantes.

REMARQUE. Les équations pour les potentiels $\left(\Delta - \frac{1}{c^2}\,\partial_t^2\right) V = -\frac{\rho}{\epsilon_0}$ et $\left(\Delta - \frac{1}{c^2}\,\partial_t^2\right)\overrightarrow{A} = -\mu_0\,\vec{j}$ (*cf.* section 7.5.2) montrent, l'opérateur **d'Alembertien** $\Delta - \frac{1}{c^2}\,\partial_t^2$ étant invariant, que $\underline{A} = (\frac{1}{c}\,V, \overrightarrow{A})$ est un quadrivecteur et que la **condition de Lorentz** $\operatorname{div}\overrightarrow{A} + \frac{1}{c^2}\,\partial_t V = 0$ ou $\underline{\partial}\cdot\underline{A} = 0$ est invariante.

4.4 PHYSIQUE QUANTIQUE ET ESPACES VECTORIELS

L'approche quantique des phénomènes, qui depuis plus de quatre-vingts ans a permis de comprendre la plupart des grandes découvertes de la physique, repose sur le calcul linéaire. En effet les **états d'un système physique** sont spécifiés par les **vecteurs normés d'un E.V. complexe** dont la dimension caractérise le nombre de degrés de liberté du système considéré. Quant aux **"opérations" sur le système** (sa préparation dans un état particulier, l'action sur lui d'une symétrie telle que translation ou rotation, une opération de mesure, son évolution, *etc.*), elles sont décrites à partir d'**applications linéaires** (opérateurs, matrices) agissant sur l'E.V. des états.

Rappelons que la physique quantique a son origine dans l'étude de l'**interaction rayonnement-matière**, résumée très schématiquement par : 1) la découverte que l'optique ondulatoire du 19ème siècle, expliquée par les ondes électromagnétiques, met en fait en jeu des particules, les photons, avec les "règles de correspondance" $E = \hbar\omega$ et $\vec{p} = \hbar\vec{k}$ pour l'énergie et la quantité de mouvement d'un photon associé à une onde plane (ces règles étant valables pour d'autres particules et concernant aussi le moment cinétique : $J_z = m\hbar$ avec m entier ou demi entier) ; 2) l'observation que la spectroscopie atomique et l'idée de photon impliquent la quantification des niveaux d'énergie, d'où la nécessité d'expliquer cette quantification ainsi que l'existence des transitions entre niveaux.

4.4.1 Cadre général ; états quantiques ; moyennes ; évolution

▷ E.V. des états quantiques

Considérons à titre d'exemple les photons associés à une onde plane électromagnétique $\begin{pmatrix} E_x \\ E_y \end{pmatrix} e^{-i(\omega t - kz)}$ ($E_{x,y} \in \mathbb{C}$). On décrit leur état de polarisation par :

$$\underline{\psi}(t) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} e^{-i\omega t} \quad \text{avec} \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = (|E_x|^2 + |E_y|^2)^{-\frac{1}{2}} \begin{pmatrix} E_x \\ E_y \end{pmatrix} .$$

L'espace (complexe) engendré par ces états est ici de dimension 2. La **préparation** d'un autre état $\underline{\varphi}$ à partir de $\underline{\psi}$ se traduit mathématiquement par l'application du projecteur $P_{\underline{\varphi}}$:

$$P_{\underline{\varphi}}\,\underline{\psi} = (\underline{\varphi}, \underline{\psi})\,\underline{\varphi} \ ;$$

le produit scalaire $(\underline{\varphi}, \underline{\psi})$ est appelé **amplitude de probabilité** de trouver l'état $\underline{\varphi}$ à partir de $\underline{\psi}$ et le carré de sa norme $|(\underline{\varphi}, \underline{\psi})|^2$, compris entre 0 et 1 (inégalité de Schwarz), est la probabilité correspondante. Par exemple $P_{\underline{\varphi}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ pour $\underline{\varphi} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ correspond à l'action d'un polariseur selon x, et $|(\underline{\varphi}, \underline{\psi})|^2 = |\alpha|^2$ est aussi le rapport des flux lumineux à la sortie et à l'entrée du polariseur.

REMARQUE. Si $\{\underline{\varphi}_i\}$ est une base orthonormée ($(\underline{\varphi}_i, \underline{\varphi}_j) = \delta_{ij}$), les probabilités

$$p_i = |(\underline{\varphi}_i, \underline{\psi})|^2$$

vérifient $\sum_i p_i = 1$ (conséquence de $\underline{\psi} = \sum_i (\underline{\varphi}_i, \underline{\psi})\,\underline{\varphi}_i$ et de $(\underline{\psi}, \underline{\psi}) = 1$)). On fera attention à ce que ce sont les amplitudes complexes $(\underline{\varphi}_i, \underline{\psi})$, et non les p_i, qui caractérisent entièrement $\underline{\psi}$ (modulo une phase globale).

Symétries. De manière générale une opération de symétrie est décrite par l'action d'un **opérateur unitaire**. Par exemple une rotation autour de l'axe z se traduit par la matrice

$$\mathbf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \exp -\frac{i}{\hbar}\theta \mathbf{J}_z \qquad \text{(définition de } \mathbf{J}_z) \ .$$

Les états propres de $\mathbf{J}_z = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ sont $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}$ (*cf.* section 4.2.3) ; les valeurs propres $\pm\hbar$ sont appelées **hélicités** du photon (moment cinétique selon la direction de propagation). Le signe + correspond à une polarisation tournant dans le sens direct puisque $\psi(t) \propto e^{-i\omega t}$ (polarisation dite gauche en optique *cf.* section 2.3.4).

▷ Moyenne quantique ; matrice densité

La **moyenne** d'un opérateur A dans un état $\underline{\psi}$, définie par

$$< A >= (\underline{\psi},\, A\underline{\psi}) \ ,$$

vaut $\sum_i a_i p_i$ (d'où le nom de moyenne) si A est diagonalisable dans une base orthonormée de vecteurs propres $\underline{\varphi}_i$ associés aux valeurs propres a_i.

EXEMPLE 1. Pour les photons la moyenne $\begin{pmatrix}\overline{\alpha} & \overline{\beta}\end{pmatrix} \mathbf{J}_z \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ de J_z dans l'état $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ vaut $\hbar\left(\left|\dfrac{\alpha - i\beta}{\sqrt{2}}\right|^2 - \left|\dfrac{\alpha + i\beta}{\sqrt{2}}\right|^2\right)$ $(= 0$ pour une polarisation rectiligne$)$; elle peut se mesurer directement à partir du couple Γ ($\Gamma =< J_z > \times$ flux de photons) qu'exerce la lumière incidente sur une lame anisotrope lorsque la polarisation sortante est rectiligne (Bethe-Harris 1935).

EXEMPLE 2. Pour $\mathbf{A} = \begin{pmatrix}0 & e^{-i\omega\tau}\\ 0 & 0\end{pmatrix}$ (non hermitienne et non diagonalisable) la moyenne quantique $< A >= \begin{pmatrix}\overline{\alpha} & \overline{\beta}\end{pmatrix}\begin{pmatrix}0 & e^{-i\omega\tau}\\ 0 & 0\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \overline{\alpha}\, e^{-i\omega\tau}\, \beta$ n'est autre que la moyenne temporelle $< \overline{E}_x(t)\, E_y(t+\tau) >$, divisée par $< |E_x|^2 + |E_y|^2 >$; elle est mesurée à partir d'expériences d'interférences (*cf.* section 2.5.3).

REMARQUES. 1) La moyenne dans un état $\underline{\psi}$ s'écrit aussi $< A >= \mathrm{tr}(AP_{\underline{\psi}})$; en effet, en choisissant une base orthonormée $\{\underline{\varphi}_i\}$ telle que $\underline{\varphi}_1 = \underline{\psi}$, on obtient : $\mathrm{tr}(AP_{\underline{\psi}}) = \sum_i (\underline{\varphi}_i, AP_{\underline{\varphi}_1}\underline{\varphi}_i) = (\underline{\varphi}_1, A\underline{\varphi}_1)$.
2) On rencontre aussi des situations expérimentales où la moyenne s'effectue par rapport à un **mélange statistique** (une superposition incohérente) d'états ; c'est le cas en optique lorsqu'on superpose des faisceaux incohérents entre eux. Si $\underline{\psi}_n$ désignent les états des photons relatifs à chaque faisceau n et p_n les "poids" relatifs des faisceaux (leurs intensités relatives qui vérifient $\sum_n p_n = 1$), alors $< A >= \sum_n p_n\, (\underline{\psi}_n,\, A\underline{\psi}_n) = \mathrm{tr}(\rho A)$ avec $\rho = \sum_n p_n P_{\underline{\psi}_n} = \sum_n p_n \begin{pmatrix}|\alpha_n|^2 & \alpha_n\,\overline{\beta}_n\\ \overline{\alpha}_n\,\beta_n & |\beta_n|^2\end{pmatrix}$. ρ est appelée **matrice densité** ; elle vérifie $\rho = \rho^\dagger$, $\mathrm{tr}\rho = 1$; elle est positive (valeurs propres $\rho_n \geq 0$) et $\mathrm{tr}\rho^2 \leq 1$; enfin $\mathrm{tr}\rho^2 = 1$ équivaut à $\rho = P_{\underline{\psi}}$ et correspond à un état pur. (*cf.* aussi section 2.5.3.)

▷ Evolution ; équation de Schrödinger

L'évolution d'un état normé est nécessairement décrite par un opérateur unitaire $\underline{\psi}(t) = U(t,t_0)\,\underline{\psi}(t_0)$, avec les propriétés $U(t',t_0) = U(t',t)\,U(t,t_0)$ et $U(t,t) = I$. Pour un intervalle de temps infinitésimal on peut poser $U(t+\epsilon,t) = I - \dfrac{i}{\hbar}\,\epsilon\, H(t)$, et on montre, comme à la section 4.2.2, que l'opérateur **hamiltonien** $H(t)$ ainsi défini est hermitien. Comme $U(t+\epsilon,t_0) = U(t+\epsilon,t)\,U(t,t_0)$, on vérifie facilement que $\dot{U}(t,t_0) = -\dfrac{i}{\hbar}\,H(t)\,U(t,t_0)$, et, par application des deux opérateurs à $\underline{\psi}(t_0)$, on obtient l'**équation de Schrödinger** :

$$i\hbar\,\dot{\underline{\psi}}(t) = H(t)\,\underline{\psi}(t) \ .$$

Si H ne dépend pas de t, sa solution est $\underline{\psi}(t) = \exp\!\big[-\dfrac{i}{\hbar}Ht\big]\,\underline{\psi}(0)$; elle vaut

$$\underline{\psi}(t) = \exp\!\big[-\frac{i}{\hbar}Et\big]\,\underline{\psi}(0)$$

si $\underline{\psi}(0)$ est état propre de H, avec la valeur propre E (**énergie de l'état**). L'évolution

laissant les états propres de H invariants, à une phase près, on les appelle **états stationnaires**. Les moyennes dans ces états de tout opérateur A indépendant du temps restent constantes lors de l'évolution ; on peut montrer que réciproquement cette propriété caractérise un état stationnaire. Une **symétrie** de H est alors un opérateur unitaire S tel que $S\underline{\psi}(t)$ obéit à la même équation que $\underline{\psi}(t)$; de $i\hbar\, S\,\dot{\underline{\psi}}(t) = S\,\left(H\underline{\psi}(t)\right) = H\,\left(S\underline{\psi}(t)\right)$, pour tout $\underline{\psi}(0)$, on déduit que S commute avec H :

$$[H, S] = 0\ .$$

REMARQUE. Pour tout A indépendant de t, une moyenne dans un état quelconque $(\underline{\psi}(t),\ A\underline{\psi}(t)) =< A >_t$ vérifie

$$\frac{\mathrm{d} < A >_t}{\mathrm{d}t} = \frac{i}{\hbar} < [H(t), A] >_t;$$

il suffit d'écrire $\underline{\psi}^\dagger(t+\epsilon)\,\mathbf{A}\,\underline{\psi}(t+\epsilon) = \underline{\psi}^\dagger(t)\left(\mathbf{I} + \frac{i}{\hbar}\epsilon\,\mathbf{H}(t)\right)\,\mathbf{A}\,\left(\mathbf{I} - \frac{i}{\hbar}\epsilon\,\mathbf{H}(t)\right)\,\underline{\psi}(t)$ (en notation matricielle), puis de développer. L'opérateur $\dot{A}$ défini par $\dot{A} = \frac{i}{\hbar}\,[H, A]$ est appelé **vitesse d'évolution** de A (*cf.* section 9.1.2 pour l'équivalent classique).

4.4.2 Dynamique des systèmes à deux états ; transitions quantiques

On considère l'exemple de la molécule NH_3 dont le plan des atomes d'hydrogène peut occuper deux positions autour de z_0 et $-z_0$ de part et d'autre de l'atome d'azote (image classique), comme une particule dans un double puits de potentiel symétrique (figures 23a et b).

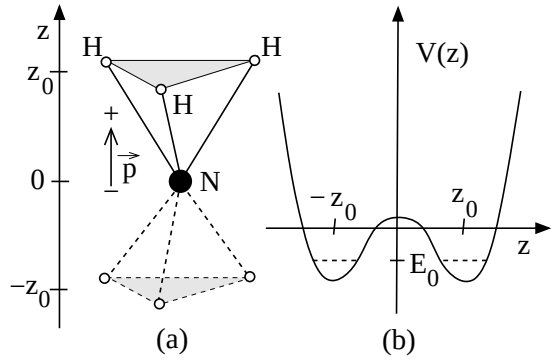


FIGURE 23

Une description quantique simplifiée, ne faisant pas appel aux fonctions d'onde, consiste à considérer que l'espace des états est engendré par deux vecteurs $\underline{\psi}_+ = \begin{pmatrix}1\\0\end{pmatrix}$ et $\underline{\psi}_- = \begin{pmatrix}0\\1\end{pmatrix}$ associés respectivement aux positions z_0 et $-z_0$; ils sont états propres de l'opérateur $\mathbf{Z} = z_0\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$ (valeurs propres $\pm z_0$) et se déduisent l'un de l'autre par l'opérateur $\mathbf{S} = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ associé à la symétrie $z \to -z$ du système.

▷ Evolution libre

Un hamiltonien obéissant à la symétrie ($\mathbf{H}\,\mathbf{S} = \mathbf{S}\,\mathbf{H}$) s'écrit alors

$$\mathbf{H}_0 = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} = E_0\,\mathbf{I} - A\,\boldsymbol{\sigma}_x$$

(avec A réel car $\mathbf{H}_0$ hermitien) ; dans la suite on suppose $A > 0$ et on pose $A = \dfrac{\hbar\omega_0}{2}$. Pour tout état initial $\underline{\psi}(0) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, l'évolution est (*cf.* $\mathbf{U}_{\hat{n}}(\theta)$ section 4.2.4) :

$$\underline{\psi}(t) = \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = e^{-\frac{i}{\hbar}t\,\mathbf{H}_0} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = e^{-\frac{i}{\hbar}\,E_0 t} \begin{pmatrix} \cos\frac{\omega_0 t}{2} & i\,\sin\frac{\omega_0 t}{2} \\ i\,\sin\frac{\omega_0 t}{2} & \cos\frac{\omega_0 t}{2} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} .$$

Elle conduit aux probabilités $|\alpha(t)|^2 = \left|\alpha\,\cos\dfrac{\omega_0 t}{2} + i\beta\,\sin\dfrac{\omega_0 t}{2}\right|^2$ d'être dans l'état $\underline{\psi}_+$ à l'instant t, et $|\beta(t)|^2 = 1 - |\alpha(t)|^2$ d'être dans l'état $\underline{\psi}_-$; par exemple $|\alpha(t)|^2 = \cos^2\dfrac{\omega_0 t}{2}$ et $|\beta(t)|^2 = \sin^2\dfrac{\omega_0 t}{2}$ si $\alpha = 1$ et $\beta = 0$; alors la position moyenne à l'instant t est $z(t) = z_0\,(|\alpha(t)|^2 - |\beta(t)|^2) = z(0)\,\cos\omega_0 t$. Les états stationnaires (vecteurs propres de $\mathbf{H}_0$, donc de $\boldsymbol{\sigma}_x$) sont l'**état fondamental** $\underline{\psi}_f = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et l'**état excité** $\underline{\psi}_e = \dfrac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ d'énergies $E_f = E_0 - A$ et $E_e = E_0 + A$. Il sont aussi états propres de $\mathbf{S}$ et pour eux $z(t) = 0$ et $\dot{z}(t) = 0$. L'hypothèse $A > 0$ garantit que $\underline{\psi}_f$ est symétrique.

▷ Polarisabilité de NH_3

La polarisation de la liaison $N - H$ se traduit, pour la molécule NH_3, par la présence d'un dipole $\vec{p}$ orienté selon l'axe z et, en présence d'un champ électrique $\mathcal{E}_z$, par une énergie supplémentaire $\mp p\mathcal{E}_z$ pour les états $\underline{\psi}_\pm$. Dans la base $\underline{\psi}_\pm$ l'évolution est donc décrite par

$$i\hbar \begin{pmatrix} \dot{\alpha} \\ \dot{\beta} \end{pmatrix} = \mathbf{H} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{avec} \quad \mathbf{H} = \begin{pmatrix} E_0 - p\mathcal{E}_z & -A \\ -A & E_0 + p\mathcal{E}_z \end{pmatrix} = E_0\,\mathbf{I} - \overrightarrow{\boldsymbol{\sigma}} \cdot \overrightarrow{V}$$

($V_x = A = V\,\sin\theta$, $V_y = 0$, $V_z = p\mathcal{E}_z = V\,\cos\theta$). Les valeurs propres sont $E_\mp = E_0 \mp \sqrt{A^2 + p^2\mathcal{E}_z^2}$ et les vecteurs propres s'écrivent (*cf.* section 4.2.3) $\begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}$ et $\begin{pmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}$. Si $\mathcal{E}_z$ est constant ces états évoluent en étant simplement multipliés par $\exp\left(-\dfrac{i}{\hbar}\,tE_\mp\right)$, et pour eux, le **moment dipolaire moyen** est :

$$< p > (t) = \pm p\,(\cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}) = \pm p\,\cos\theta \simeq \pm\frac{p^2\mathcal{E}_z}{A} \quad \text{si} \quad |p\mathcal{E}_z| \ll A\ ;$$

pour l'état fondamental il a le signe de $\mathcal{E}_z$, conformément à la description classique de la polarisabilité de NH_3, et décroît quand A augmente (interprétation physique laissée au lecteur).

▷ Transitions ; résonance ; règle d'or de Fermi

On considère un champ $\mathcal{E}_z(t)$ dépendant du temps. Dans la base $\underline{\psi}_{f,e}$ l'évolution d'un état $\underline{\psi} = c_f\,\underline{\psi}_f + c_e\,\underline{\psi}_e$ est donnée par ($c_{f,e} = \dfrac{1}{\sqrt{2}}\,(\beta \pm \alpha)$) :

$$i\hbar \begin{pmatrix} \dot{c}_f \\ \dot{c}_e \end{pmatrix} = \begin{pmatrix} E_f & p\mathcal{E}_z(t) \\ p\mathcal{E}_z(t) & E_e \end{pmatrix} \begin{pmatrix} c_f \\ c_e \end{pmatrix} \ .$$

Le changement de variables $c_{f,e}(t) = \exp\left(-\dfrac{i}{\hbar}\,E_{f,e}t\right)\,\gamma_{f,e}(t)$, qui conduit aux équations

$$i\hbar\,\dot{\gamma}_f(t) = p\,\mathcal{E}_z(t)\,e^{-i\omega_0 t}\,\gamma_e(t) \quad \text{et} \quad i\hbar\,\dot{\gamma}_e(t) = p\,\mathcal{E}_z(t)\,e^{i\omega_0 t}\,\gamma_f(t)\ ,$$

est bien adapté à l'étude des transitions entre $\underline{\psi}_e$ et $\underline{\psi}_f$ induites par $\mathcal{E}_z(t)$ car $\dot{\gamma}_f(t) = \dot{\gamma}_e(t) = 0$ pour $\mathcal{E}_z(t) = 0$. Si $\mathcal{E}_z(t)\,e^{\pm i\omega_0 t}$ est une fonction qui oscille vite, on s'attend à ce que $\dot{\gamma}_{f,e}(t) \simeq 0$, et donc à l'absence de transitions. Considérons alors le cas parfaitement monochromatique $\mathcal{E}_z(t) = \mathcal{E}_0\,e^{-i\omega t} + \overline{\mathcal{E}_0}\,e^{i\omega t}$ avec ω proche de ω_0. On obtient

$$i\hbar\,\dot{\gamma}_f(t) \simeq p\overline{\mathcal{E}}_0 e^{i(\omega-\omega_0)t}\,\gamma_e(t) \ \text{ et } \ i\hbar\,\dot{\gamma}_e(t) \simeq p\mathcal{E}_0 e^{-i(\omega-\omega_0)t}\,\gamma_f(t)$$

en ne gardant en facteur de $\gamma_{f,e}(t)$ que les coefficients à variation lente ; cette approximation, identique à la méthode des équations d'amplitude de la section 6.4.5, ne sera valide que si les solutions ainsi trouvées sont à variation lente par rapport à $e^{\pm i\omega_0 t}$. Alors $\gamma_e(t)$ satisfait l'E.D. à coefficients constants $\ddot{\gamma}_e + i(\omega-\omega_0)\,\dot{\gamma}_e + \Omega^2\,\gamma_e = 0$ avec $\Omega = \dfrac{p|\mathcal{E}_0|}{\hbar}$. Si on suppose qu'initialement l'un des états n'est pas occupé, par exemple $\gamma_e(0) = 0$ et $\gamma_f(0) = 1$, on obtient (petit calcul) :

$$|\gamma_e(t)|^2 = \Omega^2\,\frac{\sin^2\sqrt{\left(\frac{\omega-\omega_0}{2}\right)^2 + \Omega^2}\,t}{\left(\frac{\omega-\omega_0}{2}\right)^2 + \Omega^2}\ .$$

L'approximation est donc valable pour $\Omega \ll \omega_0$ et $|\omega - \omega_0| \ll \omega_0$.

Transition résonnante. Si $\omega = \omega_0$ on obtient $|\gamma_e(t)|^2 = \sin^2\Omega t$ (résultat qui découle aussi plus directement des équations $\ddot{\gamma}_{f,e} + \Omega^2\,\gamma_{f,e} = 0$). Ce basculement réversible de $\underline{\psi}_f$ vers $\underline{\psi}_e$, lorsque $\omega = \omega_0$, est appelé **résonance quantique**.

REMARQUE. Si $\omega \neq \omega_0$, et dans la mesure où les termes en $\mathcal{E}_z(t)$ dans le hamiltonien peuvent être considérés comme perturbatifs, la solution ci-dessus devient, au premier ordre du calcul de perturbation,

$$|\gamma_e(t)|^2 = \Omega^2\,F(\omega-\omega_0,t) \ \text{ avec } \ F(\omega-\omega_0,t) = \frac{\sin^2\frac{\omega-\omega_0}{2}t}{\left(\frac{\omega-\omega_0}{2}\right)^2}\ .$$

Cette expression, qui s'obtient aussi par intégration de l'E.D. pour $\gamma_e(t)$ entre 0 et t en maintenant $\gamma_f(t)$ égal à un, suppose $|\gamma_e(t)| \ll 1$. Elle reste valable pour tout ω si Ω est arbitrairement petit.

Règle d'or de Fermi. Souvent, dans la réalité, le champ $\mathcal{E}(t)$ n'est que quasi-monochromatique et l'énergie du niveau excité présente une dispersion. Ces deux phénomènes ont pour effet de distribuer la valeur de Ω^2 sur un continuum de valeurs de ω (pour le premier) et de E_e (pour le second), et donc de rendre pour chaque ω et E_e fixé la valeur correspondante de Ω^2 arbitrairement petite. Dans le premier cas, pour tenir compte de la dispersion en ω de $\mathcal{E}(t)$, il faut remplacer $\Omega^2 = \hbar^{-2}p^2|\mathcal{E}_0|^2$ par $\hbar^{-2}p^2 I_0(\omega)\,\mathrm{d}\omega$ sur l'intervalle $[\omega,\, \omega+\mathrm{d}\omega]$ ($I_0(\omega)$ densité spectrale d'intensité du champ) ; $|\gamma_e|^2$ devient $\hbar^{-2}\int p^2\, I_0(\omega)\, F(\omega-\omega_0,t)\,\mathrm{d}\omega$. Comme $F(\omega-\omega_0,t)$, dont l'intégrale sur ω vaut exactement $2\pi t$, est de plus en plus "piquée" autour de ω_0 lorsque $t\to\infty$ ($F(0,t)=t^2$ et $\Delta\omega \simeq 2\pi t^{-1}$), cette intégrale pour t grand vaut $t\, w_{ef}$ avec :

$$w_{ef} = 2\pi\hbar^{-2}p^2 I_0(\omega_0)\ .$$

w_{ef} est **la probabilité de transition par unité de temps** de $\underline{\psi}_f$ à $\underline{\psi}_e$. (En considérant le cas $\gamma_e(0)=1$ on calcule de même w_{fe} et on vérifie que $w_{fe}=w_{ef}$: **microréversibilité**). Dans le second cas, pour tenir compte de la dispersion de E_e et donc aussi de $\omega_0 = \hbar^{-1}(E_e-E_f)$, on introduit une densité $\rho_e(E)$ de niveaux excités par unité d'énergie (ce qui revient à remplacer Ω^2 par $\Omega^2\,\rho_e(E)\,\mathrm{d}E$ sur l'intervalle $[E, E+\mathrm{d}E]$) et $|\gamma_e|^2$ devient $\hbar^{-2}\,|p\,\mathcal{E}_0|^2\ \displaystyle\int \rho_e(E)\, F(\omega-\frac{E-E_f}{\hbar},t)\,\mathrm{d}E = t\, w_{\{e\}f}$ avec :

$$w_{\{e\}f} \simeq 2\pi\hbar^{-1}|p\mathcal{E}_0|^2\,\rho_e(E_f+\hbar\omega) \qquad \textbf{(règle d'or de Fermi)}\ .$$

Dans les deux cas le comportement de F montre que la transition est impossible lorsque l'égalité $E_f = E_e - \hbar\omega$ (**L.C. de l'énergie**) n'est pas satisfaite. La présence de probabilités de transition traduit le caractère irréversible de ces transitions (*cf.* section 10.3.2).

▷ Interactions avec d'autres systèmes à deux états ; loi exponentielle

Ci-dessus les transitions sont dues au terme de couplage $p\mathcal{E}_z$ présent dans le hamiltonien. Des calculs semblables conduisent aux mêmes résultats en supposant que la molécule NH_3 interagit avec un autre système $\mathcal{S}$ à deux niveaux, par exemple un **champ quantique** dont les états $\underline{\varphi}_{f,e}$ à n ou $n+1$ photons (n fixé) ont des énergies $\epsilon_{f,e}$ avec $\epsilon_e-\epsilon_f=\hbar\omega$. Si dans l'E.V. engendré par les quatre vecteurs $\underline{\psi}_{f,e}\otimes\underline{\varphi}_{f,e}$ on suppose que les transitions n'ont lieu qu'entre les deux états $\underline{\psi}_e\otimes\underline{\varphi}_f$ et $\underline{\psi}_f\otimes\underline{\varphi}_e$, ce qui est vrai si $\hbar\omega\simeq\hbar\omega_0$, on est conduit à traiter l'ensemble $NH_3-\mathcal{S}$ comme un système à deux états. Les équations d'évolution pour les composantes d'un état quelconque sur ces deux vecteurs sont alors de la forme $i\hbar\dot{c}_{fe} = (E_f+\epsilon_e)\,c_{fe} + K\,c_{ef}$ et $i\hbar\dot{c}_{ef} = (E_e+\epsilon_f)\,c_{ef} + K\,c_{fe}$, et un changement de variable semblable à celui effectué plus haut les change en $i\hbar\dot{\gamma}_{fe} = K\,e^{i(\omega-\omega_0)t}\,\gamma_{ef}$ et $i\hbar\dot{\gamma}_{ef} = K\,e^{-i(\omega-\omega_0)t}\,\gamma_{fe}$.

On peut modéliser la **loi exponentielle de désexcitation spontanée** de NH_3 en supposant que NH_3 interagit non pas avec un, mais avec une infinité de systèmes à deux niveaux dont les différences d'énergie $\hbar\omega_k$ forment un continuum autour de $\hbar\omega_0$; ce sont en pratique les modes du champ électromagnétique dans le vide, avec zéro ou un photon. On suppose que les transitions s'effectuent entre l'état $\underline{\Psi}_e$ (NH_3 dans l'état excité et les modes avec zéro photon), et les états $\underline{\Psi}_{f,k}$ (le mode k avec un photon et NH_3 et les autres modes dans leur état fondamental). La résolution exacte des équations $i\hbar\dot{\gamma}_{f,k} = K_k\,e^{i(\omega_k-\omega_0)t}\,\gamma_e$ et $i\hbar\dot{\gamma}_e = \sum_k K_k\,e^{-i(\omega_k-\omega_0)t}\,\gamma_{f,k}$ par transformée de Laplace (*cf.* section 5.3.3), avec $\gamma_e(0)=1$ et $\gamma_{f,k}(0)=0$, donne (petit calcul) : $p\,\Gamma_e(p) - 1 = -\frac{1}{\hbar^2}\sum_k K_k^2\,[p+i(\omega_k-\omega_0)]^{-1}\,\Gamma_e(p)$. Si on pose $\omega_k = \omega_0 + k\,\Delta\omega$, $K_k = K$ pour tout $k\in\mathbb{Z}$ on peut utiliser la formule exacte :

$$\textstyle\sum_{k=-\infty}^{\infty}(p+ik\,\Delta\omega)^{-1} = \frac{\pi}{\Delta\omega}\,\mathrm{cotanh}\,\frac{\pi p}{\Delta\omega}\ .$$

Dans la limite $\Delta\omega\to 0$ et $\frac{\pi K^2}{\hbar^2\,\Delta\omega}\to a$, elle donne $\Gamma_e(p) = \frac{1}{p+a}$ d'où $\gamma_e(t) = e^{-at}$. La loi exponentielle

et son irréversibilité sont donc associées à des transitions de l'état excité vers un continuum d'états ; à l'opposé, si il y a transition vers un seul état avec $\omega = \omega_0$, on obtient $\Gamma_e(p) = \frac{p}{p^2+\Omega^2}$ (avec $\Omega = \frac{K}{\hbar}$) et on retrouve la résonance $\gamma_e(t) = \cos \Omega t$.

4.4.3 Fonctions d'ondes ; E.D.P. de Schrödinger ; états gaussiens

▷ Degrés de liberté spatiaux ; translation ; percussion

Pour décrire ces degrés de liberté, en se limitant à une particule et à une seule dimension d'espace, il faut introduire l'E.V., de dimension infinie, des fonctions de z muni du produit scalaire $(g, f) = \int \overline{g}(z)\, f(z)\, \mathrm{d}z$. La composante spatiale d'un état est alors représentée par une fonction normée $\psi(z,t)$ appelée fonction d'onde ou encore "paquet d'ondes" car, en vertu de l'analyse de Fourier, elle peut s'écrire comme une "somme" $\int \widehat{\psi}(k,t)\, e^{ikz}\, \mathrm{d}z$ d'ondes planes. Dans cet espace l'opérateur **quantité de mouvement** est $P = \dfrac{\hbar}{i}\dfrac{\partial}{\partial z}$ (noter que $Pf \simeq p_0 f$ si f est la fonction d'onde d'une particule de quantité de mouvement p_0 quasi fixée, c.a.d. si f est "proche" de $\exp\Big[\dfrac{i}{\hbar} p_0 z\Big]$) ; celui de **position** Z est la multiplication par z ($Zf \simeq z_0 f$ si f est "piquée" en z_0, *cf.* l'interprétation de $|f(z)|^2$ ci-dessous). L'opérateur $\exp\Big[-\dfrac{i}{\hbar}\, aP\Big]$, qui change $f(z)$ en $f(z-a)$ (*cf.* développement de Taylor section 1.1.3), correspond à une **translation** (figure 24). On vérifie qu'il change $< Z >= \int \overline{f}(z)\, z\, f(z)\, \mathrm{d}z$ en $< Z > +a$ et laisse inchangé $< P >= \displaystyle\int \overline{f}(z)\, \frac{\hbar}{i}\, f'(z)\, \mathrm{d}z$. L'opérateur $\exp\Big[\dfrac{i}{\hbar}\, pZ\Big]$, qui multiplie $f(z)$ par $\exp \dfrac{ipz}{\hbar}$, ne change pas $< Z >$ mais change $< P >$ en $< P > +p$; il correspond à une "**percussion**" de l'état (figure 24). Z et P vérifient :

$$Z = Z^\dagger \quad , \quad P = P^\dagger \quad \text{et} \quad [Z, P] = i\hbar\ .$$

DÉMONSTRATION. L'adjoint d'un opérateur étant défini par $(g, Af) = (A^\dagger g, f)$ (comme pour les matrices), le résultat est trivial pour Z et se déduit d'une intégration par parties de $\displaystyle\int \overline{g}(z)\, \frac{\hbar}{i}\, f'(z)\, \mathrm{d}z$ pour P. Enfin la relation de commutation $[Z, P] = i\hbar$ résulte de $\left(z\dfrac{\partial}{\partial z}\right) f(z) - \dfrac{\partial}{\partial z}\,(z\, f(z)) = -f(z)$.

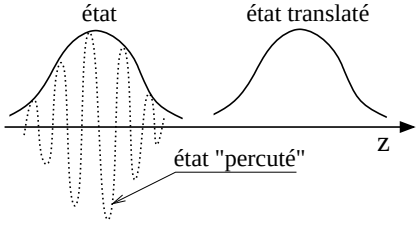


FIGURE 24

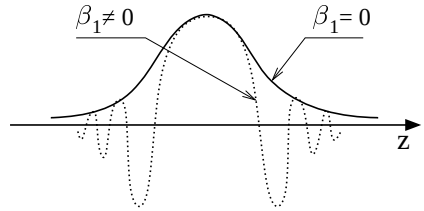


FIGURE 25

L'inégalité de Schwarz (*cf.* section 4.1.2) pour les états $(Z - < Z >)\, f$ et $(P - < P >)\, f$ conduit, comme pour les signaux, à l'**inégalité de Heisenberg** :

$$\delta z\ \delta p \geq \frac{\hbar}{2} \quad \text{avec} \quad (\delta z^2) =< (Z - < Z >)^2 > \quad \text{et} \quad (\delta p^2) =< (P - < P >)^2 > \ .$$

▷ E.D.P. de Schrödinger ; lien avec la mécanique classique

On se limite ici au cas de particules galiléennes de masse m. Pour des particules libres on pose $H = \dfrac{P^2}{2m}$ (noter que $Hf \simeq \dfrac{p_0^2}{2m} f$ pour f "proche" de $\exp\Big[\dfrac{i}{\hbar} p_0 z\Big]$). Si les particules ont classiquement une énergie potentielle $V(z)$, un choix fréquent (discuté ci-dessous) est $H = \dfrac{P^2}{2m} + V(Z)$. L'équation $i\hbar\, \underline{\dot{\psi}}(t) = H\, \underline{\psi}(t)$ devient alors l'**E.D.P. de Schrödinger** (équation d'onde) :

$$i\hbar\, \frac{\partial \psi(z,t)}{\partial t} = -\frac{\hbar^2}{2m}\, \frac{\partial^2 \psi(z,t)}{\partial z^2} + V(z)\, \psi(z,t)\ .$$

Elle conduit à la loi de conservation locale :

$$\frac{\partial \rho(z,t)}{\partial t} + \frac{\partial j(z,t)}{\partial z} = 0 \quad \text{avec} \quad \rho = |\psi|^2 \ \text{ et } \ j = -i\frac{\hbar}{2m}\left(\overline{\psi}\, \frac{\partial \psi}{\partial z} - \psi\, \frac{\partial \overline{\psi}}{\partial z}\right)\ .$$

$\rho(z,t)$ est interprété comme une **densité de probabilité** de présence de la particule en z à l'instant t et $j(z,t)$ comme un **courant de probabilité** ; si ψ est proche d'une onde plane $\exp\Big[\dfrac{i}{\hbar} pz\Big]$, alors $j \simeq \rho v$ avec $v = \dfrac{p}{m}$, expression classique de la vitesse, ou vitesse de groupe du paquet d'ondes (cf section 8.2.1).

REMARQUE. Les moyennes de Z et P obéissent aux équations $\dfrac{\mathrm{d}}{\mathrm{d}t} < Z >= \dfrac{< P >}{m}$ et $\dfrac{\mathrm{d}}{\mathrm{d}t} < P >= - < V'(Z) >$ (par exemple $\dfrac{\mathrm{d}}{\mathrm{d}t} < P >= \dfrac{i}{\hbar}\, (\psi, [H,P]\psi)$ avec $\dfrac{i}{\hbar}\, [H,P]\psi = V\, \partial_z \psi - \partial_z (V\psi) = -V'\psi$). Ce résultat, qui ressemble aux équations de la **mécanique classique** (sauf que $< V'(Z) >= V'(< Z >)$ uniquement pour V quadratique), justifie en partie le choix de H fait plus haut. L'introduction des interactions électromagnétiques à partir d'un principe de jauge est faite à la section 9.2.2.

▷ Etats gaussiens ; états cohérents ; lien avec l'espace de phase

La famille de fonctions

$$\psi(z) \propto \exp\left[\frac{i}{\hbar}\left(p_0 z + \frac{\beta}{2}\,(z - z_0)^2\right)\right]\ ,$$

avec p_0 et z_0 réels et $\beta = \beta_1 + i\beta_2$ ($\beta_2 > 0$) se prête à des calculs explicites et est stable vis-à-vis des évolutions associées à des hamiltoniens au plus quadratiques dans le couple (Z,P). $\psi(z)$ qui vérifie $|\psi|^2 \propto \exp\left[-\dfrac{\beta_2}{\hbar}\,(z - z_0)^2\right]$ est représentée (pour $p_0 = 0$) sur la figure 25.

Dispersion en z et p. On déduit des propriétés des gaussiennes réelles que $< Z >= z_0$ et que $(\delta z)^2 =< (Z - z_0)^2 >= \hbar(2\beta_2)^{-1}$. Comme $(P\psi)(z) = (p_0 + \beta(z - z_0))\ \psi(z)$ on a aussi $< P >= p_0$ et $(\delta p)^2 =< (P - p_0)^2 >= \hbar|\beta|^2(2\beta_2)^{-1}$; en particulier $\delta z\, \delta p = \dfrac{\hbar}{2}$ si $\beta_1 = 0$. Enfin, il résulte de $[P - p_0, Z - z_0] = -i\hbar$ que :

$$\frac{1}{2} < (P-p_0)(Z-z_0)+(Z-z_0)(P-p_0) >=< (Z-z_0)(P-p_0) > -i\frac{\hbar}{2} = \hbar\beta_1(2\beta_2)^{-1} \ .$$

Si on assimile ces valeurs moyennes aux corrélations Γ de deux v.a. gaussiennes Z et P (*cf.* section 10.2.1), leur densité de probabilité conjointe est proportionnelle à

$$\exp - \left[\frac{1}{\hbar\beta_2} \left(|\beta|^2 (z-z_0)^2 - 2\beta_1 (z-z_0)(p-p_0) + (p-p_0)^2 \right) \right] \ ,$$

résultat évident si $\beta_1 = 0$. En écrivant que le crochet $[\dots]$ est par exemple plus petit que 2 on peut associer à ψ une tache elliptique dans l'**espace de phase** (z,p) assimilable au support de la densité de probabilité (figure 26). Cette tache, d'aire h indépendante de β (petit calcul ; cf section 4.2.3), est inclinée par rapport aux axes (z,p) si $\beta_1 \neq 0$ et inscrite dans un rectangle de côtés $4\,\delta z$, $4\,\delta p$ (propriété facile à vérifier si $\beta_1 = 0$).

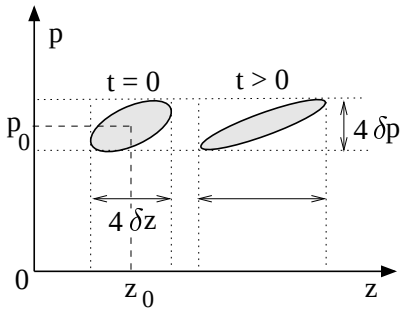


FIGURE 26

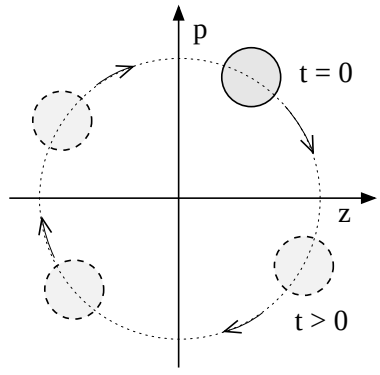


FIGURE 27

Evolution. Si $H = \dfrac{P^2}{2m} + \dfrac{1}{2}\, m\omega^2 Z^2$ (**oscillateur harmonique**), une évolution infinitésimale $\exp\left[-\dfrac{i}{\hbar}\,\epsilon H\right] = 1 - \dfrac{i}{\hbar}\,\epsilon H$ donne :

$$\begin{aligned} \psi(z) \quad &\to \psi(z) \left(1 - i\tfrac{\epsilon}{2m\hbar} \left((p_0 + \beta(z-z_0))^2 - i\beta\hbar + m^2\omega^2 z^2 \right)\right) \\ &\propto \psi(z) \exp\left[-i\tfrac{\epsilon}{2m\hbar} \left((p_0 + \beta(z-z_0))^2 + m^2\omega^2 z^2 \right)\right] \ . \end{aligned}$$

Elle conserve donc le caractère gaussien. Par identification on voit que, dans le temps ϵ, $\delta\beta = -\epsilon(m\omega^2 + m^{-1}\beta^2)$ et (petit calcul) $\delta p_0 = -\epsilon m\omega^2 z_0$ et $\delta z_0 = \epsilon m^{-1} p_0$. On en déduit les équations d'évolution :

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = -(m\omega^2 + m^{-1}\beta^2) \quad , \quad \frac{\mathrm{d}p_0}{\mathrm{d}t} = -m\omega^2 z_0 \ \text{ et } \ \frac{\mathrm{d}z_0}{\mathrm{d}t} = m^{-1} p_0 \ .$$

Les deux dernières sont attendues (*cf.* remarque ci-dessus) ; la première est identique à celle vérifiée par $\dfrac{p_0}{z_0}$, ce qui permet de la résoudre facilement.

EXEMPLES. 1) $\omega = 0$ (**particule libre**)

$$\begin{pmatrix} z_0(t) \\ p_0(t) \end{pmatrix} = \begin{pmatrix} 1 & \frac{t}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z_0(0) \\ p_0(0) \end{pmatrix} \quad \text{et donc} \quad \beta(t) = \beta(0) \left(1 + \frac{t}{m}\, \beta(0) \right)^{-1} \; ;$$

cette expression permet de discuter l'étalement en z d'un paquet d'ondes (figure 26). On vérifie que :

$$(\delta p)^2(t) = (\delta p)^2(0) \quad ; \quad |\beta(t)|^2\, (\delta z)^2(t) = |\beta(0)|^2\, (\delta z)^2(0) \; .$$

2) Si $\beta(0) = im\omega$ on a $\beta(t) = \beta(0)$; un tel état, dont la tache dans le plan (z, p) ne se déforme pas au cours du temps, est appelé **état cohérent** (figure 27 où $m = \omega = 1$).

REMARQUE. Dans le cas général :

$$\begin{pmatrix} z_0(t) \\ p_0(t) \end{pmatrix} = \begin{pmatrix} \cos \omega t & (m\omega)^{-1} \sin \omega t \\ -m\omega \sin \omega t & \cos \omega t \end{pmatrix} \begin{pmatrix} z_0(0) \\ p_0(0) \end{pmatrix} \; .$$

L'évolution des états gaussiens est donc entièrement déterminée par des matrices de transfert de déterminant 1 reliant $\begin{pmatrix} z_0(t) \\ p_0(t) \end{pmatrix}$ à $\begin{pmatrix} z_0(0) \\ p_0(0) \end{pmatrix}$. Le lecteur fera le lien avec l'**optique matricielle**, z et p jouant le rôle des coordonnées x et p d'un rayon et t celui de la coordonnée z sur l'axe optique. L'analogue des ces états gaussiens sont, en optique ondulatoire, les **faisceaux gaussiens** (ondes sphériques si $\beta_2 \simeq 0$).

4.4.4 Oscillateur harmonique ; champ électromagnétique

▷ Oscillateur harmonique ; opérateurs d'annihilation et de création

Si dans l'hamiltonien

$$H = \frac{P^2}{2m} + \frac{1}{2}\, m\omega^2 Z^2 \; ,$$

on remplace $\hbar$, m et ω par 1, ceci revient à choisir comme unité de temps, de longueur et d'énergie respectivement ω^{-1}, $\left(\dfrac{\hbar}{m\omega} \right)^{\frac{1}{2}}$ et $\hbar\omega$. On vérifie alors facilement qu'en introduisant les opérateurs

$$a = \frac{Z + iP}{\sqrt{2}} \quad , \quad a^\dagger = \frac{Z - iP}{\sqrt{2}} \quad ([a, a^\dagger] = 1 \text{ car } [Z, P] = i) \; ,$$

on obtient l'expression de H et les relations de commutation :

$$H = a^\dagger a + \frac{1}{2} \quad , \quad [H, a] = -a \quad , \quad [H, a^\dagger] = a^\dagger \; .$$

On en déduit (démonstration ci-dessous) que les **états propres** $\underline{\psi}_n$ de H forment une suite d'états orthogonaux reliés les uns aux autres par a et $a^\dagger$ et que les **niveaux d'énergie** sont quantifiés et régulièrement espacés :

$$H\, \underline{\psi}_n = \left(n + \frac{1}{2} \right) \underline{\psi}_n \quad \text{avec } a^\dagger\, \underline{\psi}_n = \sqrt{n+1}\, \underline{\psi}_{n+1} \text{ et } a\, \underline{\psi}_n = \sqrt{n}\, \underline{\psi}_{n-1} \quad (n = 0, 1, 2 \cdots) \; .$$

($a\, \underline{\psi}_0 = 0$ pour l'état fondamental $\underline{\psi}_0$.) L'action sur $\underline{\psi}_n$ de a et $a^\dagger$ faisant disparaître ou apparaître un quantum d'énergie, a et $a^\dagger$ sont appelés **opérateurs d'annihilation et**

de création. Comme $< \dot{a} >= i < [H,a] >= -i < a >$, leur moyenne dans un état $\underline{\psi}(t)$ vérifie (en rétablissant ω) :

$$< a > (t) = e^{-i\omega t} < a > (0) \quad \text{et} \quad < a^\dagger > (t) = e^{i\omega t} < a^\dagger > (0) \ .$$

DÉMONSTRATION. Soit $\underline{\psi}$ un état tel que $H\underline{\psi} = E\underline{\psi}$, alors $\underline{\phi} = a\underline{\psi}$ vérifie $H\underline{\phi} = (E-1)\,\underline{\phi}$ (car $Ha\underline{\psi} = (aH - a)\,\underline{\psi}$) et $(\underline{\phi},\underline{\phi}) = (\underline{\psi}, a^\dagger a\underline{\psi}) = E - \frac{1}{2}$. Comme $(\underline{\phi},\underline{\phi})$ est toujours positif il en résulte l'existence d'un état d'énergie minimale $\underline{\psi}_0$ tel que $E_0 = \frac{1}{2}$ et $a\underline{\psi}_0 = 0$ (unique puisque $(z+\partial_z)\,\psi_0(z) = 0$ définit la fonction d'onde $\psi_0(z) = (\pi)^{-\frac{1}{4}}\ \exp -\frac{z^2}{2}$). Comme $H\,(a^\dagger\underline{\psi}) = (E+1)\,a^\dagger\underline{\psi}$ (et $\parallel a^\dagger\underline{\psi}\parallel^2 = E+\frac{1}{2}$), les états $\underline{\psi}_n$ se déduisent de $\underline{\psi}_0$ par action de $a^\dagger$; l'introduction du facteur $\sqrt{n+1}$ dans le passage de $\underline{\psi}_n$ à $\underline{\psi}_{n+1}$ assure que $\parallel \underline{\psi}_{n+1} \parallel^2 = 1$ si $\parallel \underline{\psi}_n \parallel^2 = 1$.

REMARQUE. **Fonctions propres**. De $\underline{\psi}_n = (n!)^{-\frac{1}{2}}\,a^{\dagger n}\,\underline{\psi}_0$, on déduit : $\psi_n(z) = (\sqrt{\pi}\,2^n\,n!)^{-\frac{1}{2}} \times \big((z-\partial_z)^n\,e^{-\frac{z^2}{2}}\big)$. $(z-\partial_z)^n\,e^{-\frac{z^2}{2}}$ s'écrit aussi $H_n(z)\,e^{-\frac{z^2}{2}}$ où $H_n(z)$ est le n$^{\text{ème}}$ **polynôme de Hermite** ($H_0(z) = 1,\ H_1(z) = z,\ H_2(z) = 4z^2 - 2,\ H_3(z) = 8z^3 - 12z\ \cdots$; *cf*. figure 28). Si on rétablit $\hbar$, ω et m on obtient :

$$\psi_n(z) = (\sqrt{\pi}\,2^n\,n!)^{-\frac{1}{2}}\ \big(\tfrac{m\omega}{\hbar}\big)^{\frac{1}{4}}\ H_n\left(\sqrt{\tfrac{m\omega}{\hbar}}\,z\right)\ e^{-\frac{1}{2}\,\frac{m\omega}{\hbar}\,z^2}$$

(en n'oubliant pas que la normalisation de $\psi_n(z)$ entraîne que sa dimension est $[L]^{-\frac{1}{2}}$).

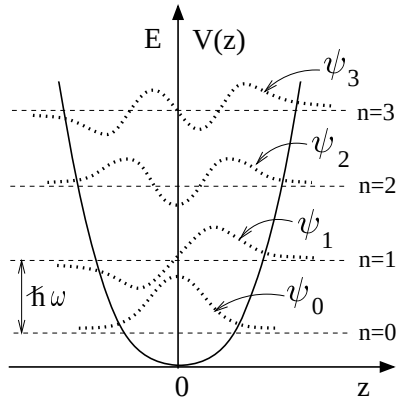


FIGURE 28

▷ Etats cohérents et limite classique

Ces états introduits plus haut par $(P\psi)(z) = (p_0 + \beta(z - z_0))\,\psi(z)$, avec $\beta = i$, sont les états propres de a. Si on les repère par leur valeur propre α, on a

$$a\,\underline{\psi}_\alpha = \alpha\,\underline{\psi}_\alpha \quad \text{avec} \quad \alpha = \frac{1}{\sqrt{2}}\,(z_0 + ip_0)\ ,$$

où z_0 et p_0 sont les valeurs moyennes de Z et P. On vérifie :

$(\underline{\psi}_\alpha, a\underline{\psi}_\alpha) =< a >= \alpha \quad , \quad (\underline{\psi}_\alpha, a^\dagger\underline{\psi}_\alpha) =< a^\dagger >= \overline{\alpha}$ et $(\underline{\psi}_\alpha, a^\dagger a\underline{\psi}_\alpha) =< a^\dagger a >= |\alpha|^2$.

Dans la base $\underline{\psi}_n$ la relation $a\,\underline{\psi}_\alpha = \alpha\,\underline{\psi}_\alpha$ et la condition de normalisation conduisent à :

$$\underline{\psi}_\alpha = \exp -\frac{|\alpha|^2}{2}\ \sum_{n=0}^{\infty}(n!)^{-\frac{1}{2}}\,\alpha^n\,\underline{\psi}_n\ ;$$

en particulier $\underline{\psi}_{\alpha=0} \equiv \underline{\psi}_{n=0}$ (état fondamental).
Pour $|\alpha| \gg 1$, l'oscillateur quantique dans un état cohérent devient assimilable à un **oscillateur classique**. En effet, dans un tel état les dispersions deviennent négligeables car $z_0 \gg \delta z = 2^{-\frac{1}{2}}$ et $p_0 \gg \delta p = 2^{-\frac{1}{2}}$ (*cf.* états gaussiens) ; l'énergie E est elle aussi quasi fixée car la loi de probabilité

$$p_n = |(\underline{\psi}_n, \underline{\psi}_\alpha)|^2 = \exp\left[-|\alpha|^2\right] \frac{|\alpha|^{2n}}{n!}$$

pour les valeurs de n dans un état $\underline{\psi}_\alpha$ (loi de Poisson, *cf.* section 10.2.2) donne une moyenne $< n >= |\alpha|^2$ et une dispersion $\sigma = |\alpha|$, d'où $\dfrac{\delta E}{E} \sim \dfrac{\sigma}{< n >} = |\alpha|^{-1} \simeq 0$. Enfin l'évolution des états $\underline{\psi}_n$ entraîne qu'un état $\underline{\psi}_{\alpha(0)}$ évolue (à une phase globale près) en $\underline{\psi}_{\alpha(t)}$ avec $\alpha(t) = \alpha(0)\, e^{-i\omega t}$; si on pose $\alpha(t) = \dfrac{1}{\sqrt{2}}\,(z(t) + ip(t))$ l'évolution donnant $\alpha(t)$ conduit à l'évolution classique pour les variables $z(t)$ et $p(t)$.

▷ Champ électromagnétique ; émissions induite et spontanée

En physique classique, ce champ défini par son action $\overrightarrow{F} = q\,(\overrightarrow{E} + \vec{v} \wedge \overrightarrow{B})$ sur une particule chargée, est une grandeur réelle. Par exemple $\overrightarrow{E} \propto \cos(\omega t - k_x x)\,\hat{z}$ pour un mode propagatif selon x et polarisé rectilignement selon z. (On rappelle que dans une boite cubique de volume $V = L^3$, un mode est défini par $(k_x, k_y, k_z) = \frac{2\pi}{L}\,(m, n, p)$ avec $m, n, p \in \mathbb{Z}$). Quand ce champ intervient dans le hamiltonien de NH_3, on a vu que les termes en $e^{i\omega t}$ et $e^{-i\omega t}$ jouent des rôles différents en induisant respectivement des transitions $\underline{\psi}_e \to \underline{\psi}_f$ et $\underline{\psi}_f \to \underline{\psi}_e$, transitions dont on sait qu'elles s'accompagnent de l'émission et de l'absorption d'un photon. Cette remarque, et aussi le fait qu'un mode est l'analogue d'un oscillateur, conduisent à remplacer la fonction classique $\mathcal{E}_z(t) = \mathcal{E}_0 e^{-i\omega t} + \overline{\mathcal{E}_0} e^{i\omega t}$ par le **champ quantique**

$$E_z = C\,(a + a^\dagger) \qquad \text{avec} \quad C = \sqrt{\frac{\hbar\omega}{2V\epsilon_0}}\ ,$$

opérateur dont la moyenne est $< E_z > (t) = C\,(< a > (0)\, e^{-i\omega t} + < a^\dagger > (0)\, e^{i\omega t})$. a et $a^\dagger$ sont les opérateurs sans dimension d'annihilation (absorption) et de création (émission) d'un photon du mode considéré et C, dont la valeur est justifiée ci-dessous, assure l'homogénéité dimensionnelle. Les états propres $\underline{\psi}_n$ de l'oscillateur désignent maintenant les états à n photons. Cette expression pour E_z permet de montrer que les probabilités de transition "atomiques" de $\underline{\psi}_e$ vers $\underline{\psi}_f$ et de $\underline{\psi}_f$ vers $\underline{\psi}_e$ en présence de n photons sont respectivement proportionnelles à $n+1$ (**émissions induite et spontanée**) et à n (**absorption**) (démonstration 1 et application à la section 10.3.2). On retrouve le **champ classique**, avec une amplitude et une phase bien définies, lorsque les états du mode sont des états cohérents à grand nombre moyen de photons (démonstration 2).

DÉMONSTRATION 1. Dans la transition $\underline{\psi}_e \to \underline{\psi}_f$ (resp. $\underline{\psi}_f \to \underline{\psi}_e$) il y a passage de n à $n+1$ photons (resp. de n à $n-1$). Dans la règle d'or de Fermi la quantité $|\mathcal{E}_0|^2$ doit alors être remplacée par $|(\underline{\psi}_{n+1}, E_z\, \underline{\psi}_n)|^2 = C^2(n+1)$ (resp. $|(\underline{\psi}_{n-1}, E_z\, \underline{\psi}_n)|^2 = C^2 n$).

DÉMONSTRATION 2. Dans un état $\underline{\psi}_\alpha$ on a $< E_z >= C\,(\alpha\, e^{-i\omega t} + \overline{\alpha}\, e^{i\omega t})$ et (petit calcul) $< E_z^2 >= < E_z >^2 + C^2$. Les fluctuations de E_z deviennent donc négligeables si $|\alpha|^2 \gg 1$ (par exemple $|\alpha|^2 > 10^{15}$

pour un faisceau laser), et E_z est alors assimilable à sa valeur moyenne. L'énergie électromagnétique classique contenue dans un volume V, qui est la moyenne temporelle de deux fois $\frac{1}{2}\,\epsilon_0 E^2 V$ (facteur 2 pour tenir compte de l'énergie magnétique), vaut $\epsilon_0 2C^2|\alpha|^2 V = |\alpha|^2\hbar\omega$, ce qui justifie la valeur de C. On notera que dans un état $\underline{\psi}_n$, $< E_z^2 >= 2C\,(n+\frac{1}{2})$ mais que $< E_z >= 0$; on interprète ce résultat en associant à un tel état un champ classique dont la phase est indéterminée (équirépartie sur $[0, 2\pi]$). C'est le cas aussi de la phase du champ associé à une lumière thermique qui correspond à un mélange statistique d'états $\underline{\psi}_n$.

4.4.5 E.D.P. relativistes de Klein-Gordon, Weyl, Dirac ; champs quantiques ; particules et antiparticules

Dans cette section on pose $\hbar = c = 1$ pour simplifier l'écriture des formules, et on se limite à la description de particules libres pour lesquelles l'énergie E positive et la quantité de mouvement $\overrightarrow{p}$ vérifient (*cf.* section 4.3.3) :

$$E = (p^2 + m^2)^{\frac{1}{2}} \quad ; \quad E = |\overrightarrow{p}| \text{ si } \; m = 0 \; .$$

Comme la relation entre E et $\overrightarrow{p}$ est non polynomiale, contrairement au cas galiléen $E = p^2/2m$, il n'y a pas d'E.D.P. relativiste ayant pour seules solutions des fonctions d'onde $\exp -i(Et - \overrightarrow{p}\cdot\overrightarrow{r})$. Les E.D.P. admettant ces solutions ont aussi des solutions $\exp i(Et - \overrightarrow{p}\cdot\overrightarrow{r})$ historiquement dites d'énergie négative et considérées alors à tort comme des fonctions d'onde d'antiparticules.

▷ E.D.P. de Klein-Gordon ; champs classiques et quantiques ; particules et antiparticules

De la même façon que l'E.D.P. de Schrödinger est associée à la relation $E = p^2/2m$ (*cf.* section 4.4.3), l'E.D.P. de **Klein-Gordon** est associée à la relation relativiste $E^2 = p^2 + m^2$. Elle s'écrit

$$(\Delta - \partial_t^2)\,\psi = m^2\psi \quad (\psi \text{ scalaire})$$

et admet, suivant que le champ est réel ou complexe, des solutions ondes planes :

$$\psi_{\overrightarrow{p}} = \alpha_{\overrightarrow{p}} e^{-i\underline{p}\cdot\underline{x}} + \overline{\alpha}_{\overrightarrow{p}} e^{i\underline{p}\cdot\underline{x}} \quad \text{ou} \quad \psi_{\overrightarrow{p}} = \alpha_{\overrightarrow{p}} e^{-i\underline{p}\cdot\underline{x}} + \overline{\beta}_{\overrightarrow{p}} e^{i\underline{p}\cdot\underline{x}}$$

(avec $\underline{p}\cdot\underline{x} = Et - \overrightarrow{p}\cdot\overrightarrow{r}$). Tant que le champ ψ réel est considéré comme classique, les amplitudes $\alpha_{\overrightarrow{p}}$ et $\overline{\alpha}_{\overrightarrow{p}}$ sont des nombres, mais si le champ est quantique ces nombres, qui sont respectivement en facteur de e^{-iEt} et e^{iEt} deviennent, comme on l'a justifié pour le champ électromagnétique, des opérateurs proportionnels aux opérateurs d'annihilation $a_{\overrightarrow{p}}$ et de création $a^\dagger_{\overrightarrow{p}}$, qui concernent ici des particules de masse m et de quantité de mouvement $\overrightarrow{p}$. On peut aussi considérer que le champ classique correspond à une somme (discrète ou continue) de modes d'oscillateurs caractérisés chacun par une valeur de $\overrightarrow{p}$ (*cf.* sections 8.1-2) ; le champ quantique est alors obtenu par quantification de ces oscillateurs.

Ce qu'apporte de nouveau un **champ complexe**, c'est l'opération de conjugaison faisant passer de $\psi_{\overrightarrow{p}}$ à :

$$\overline{\psi}_{\overrightarrow{p}} = \overline{\alpha}_{\overrightarrow{p}} e^{i\underline{p}\cdot\underline{x}} + \beta_{\overrightarrow{p}} e^{-i\underline{p}\cdot\underline{x}} \; .$$

Si le champ est quantifié $\alpha_{\overrightarrow{p}}$, $\overline{\alpha}_{\overrightarrow{p}}$ et $\beta_{\overrightarrow{p}}$, $\overline{\beta}_{\overrightarrow{p}}$ deviennent à un facteur près des opérateurs d'annihilation et de création $a_{\overrightarrow{p}}$, $a^\dagger_{\overrightarrow{p}}$ et $b_{\overrightarrow{p}}$, $b^\dagger_{\overrightarrow{p}}$, associés respectivement à la **particule** et à son **antiparticule** (l'appellation "particule" étant conventionnelle puisque ψ et $\overline{\psi}$ jouent des rôles semblables).

▷ E.D.P. de Weyl ; particules de masse nulle et d'hélicité $\pm\frac{1}{2}$

L'E.D.P. de Klein-Gordon est du second ordre en temps. Les E.D.P. invariantes relativistes du premier ordre les plus simples sont les E.D.P. conjugées (la notion de conjugaison étant celle introduite à la section 4.2.4) :

$$\boldsymbol{\partial}\underline{\widetilde{\psi}} = 0 \quad , \quad \widetilde{\boldsymbol{\partial}}\underline{\psi} = 0 \qquad (\boldsymbol{\partial} = \mathbf{I}\partial_t - \overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{\nabla}\ ;\ \widetilde{\boldsymbol{\partial}} = \mathbf{I}\partial_t + \overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{\nabla}) \ .$$

Elles concernent des **spineurs** et ont pour solutions onde plane

$$\underline{\widetilde{\psi}}_{\overrightarrow{p}} = \left(\alpha_{\overrightarrow{p}}\, e^{-i\underline{p}\cdot\underline{x}} + \overline{\beta}_{\overrightarrow{p}}\, e^{i\underline{p}\cdot\underline{x}}\right)\underline{m}_-(\hat{p}) \quad , \quad \underline{\psi}_{\overrightarrow{p}} = \left(\beta_{\overrightarrow{p}}\, e^{-i\underline{p}\cdot\underline{x}} + \overline{\alpha}_{\overrightarrow{p}}\, e^{i\underline{p}\cdot\underline{x}}\right)\underline{m}_+(\hat{p}) \ ,$$

avec $E^2 - \overrightarrow{p}^2 = 0$ et $(\overrightarrow{\boldsymbol{\sigma}}\cdot\hat{p})\,\underline{m}_\pm(\hat{p}) = \pm\underline{m}_\pm(\hat{p})$. Si on se limite au terme en $e^{-ip\cdot x}$, on peut considérer que $\underline{\widetilde{\psi}}_{\overrightarrow{p}}$ décrit une particule de **masse nulle** et d'**hélicité** $-1/2$, par exemple un **neutrino** ; $\underline{\psi}_{\overrightarrow{p}}$ décrit alors un **antineutrino** (masse nulle, hélicité $1/2$). Quand le champ est quantifié, $\alpha_{\overrightarrow{p}}$, $\overline{\alpha}_{\overrightarrow{p}}$ et $\beta_{\overrightarrow{p}}$, $\overline{\beta}_{\overrightarrow{p}}$ deviennent proportionnels à des opérateurs d'annihilation et de création de particule (neutrino) et antiparticule (antineutrino).

JUSTIFICATION. L'invariance résulte des lois de transformation de $\underline{\psi}$ et $\underline{\widetilde{\psi}}$ et de ce que $\boldsymbol{\partial}$ et $\widetilde{\boldsymbol{\partial}}$ se transforment comme $\mathbf{X}$ et $\widetilde{\mathbf{X}}$ puisque $(\partial_t, -\overrightarrow{\nabla})$ est un quadrivecteur ; par exemple $\boldsymbol{\partial}'\underline{\widetilde{\psi}}' = (\mathbf{M}\boldsymbol{\partial}\mathbf{M}^\dagger)(\mathbf{M}^{\dagger-1}\underline{\widetilde{\psi}}) = 0$ équivaut à $\boldsymbol{\partial}\underline{\widetilde{\psi}} = 0$. La conjugaison des équations se déduit de $\widetilde{\boldsymbol{\partial}} = \boldsymbol{\epsilon}\overline{\boldsymbol{\partial}}\boldsymbol{\epsilon}^{-1}$; par exemple si $\boldsymbol{\partial}\underline{\widetilde{\psi}} = 0$, on a $\boldsymbol{\epsilon}\overline{\boldsymbol{\partial}}\boldsymbol{\epsilon}^{-1}\boldsymbol{\epsilon}\overline{\underline{\widetilde{\psi}}} = 0$ et donc aussi $\widetilde{\boldsymbol{\partial}}(\boldsymbol{\epsilon}\overline{\underline{\widetilde{\psi}}}) = 0$. Si on cherche les solutions $\exp \pm i\underline{p}\cdot\underline{x}$ de $\boldsymbol{\partial}\underline{\widetilde{\psi}} = 0$ par exemple, on est conduit à $(E\mathbf{I} + \overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{p})\underline{\widetilde{\psi}} = 0$, ce qui implique $\det(E\mathbf{I} + \overrightarrow{\boldsymbol{\sigma}}\cdot\overrightarrow{p}) = E^2 - \overrightarrow{p}^2 = 0$ (masse nulle) et $(\overrightarrow{\boldsymbol{\sigma}}\cdot\hat{p})\underline{\widetilde{\psi}} = -\underline{\widetilde{\psi}}$, donc $\underline{\widetilde{\psi}}$ proportionnel à $\underline{m}_-(\hat{p})$ (*cf.* section 4.2.4). Dans une rotation d'axe $\hat{p}$ et d'angle θ, $\underline{m}_-(\hat{p})$ est multiplié par $e^{i\frac{\theta}{2}}$ (action de $\mathbf{U}_{\hat{p}}(\theta)$). Si on se rappelle (*cf.* sections 3.4.2, 4.2.5 et 4.4.1) qu'une fonction d'onde correspondant à un moment cinétique $J_z = m$ selon z est multipliée par $e^{-im\theta}$ dans un rotation d'axe $\hat{z}$ et d'angle θ, on en déduit que l'onde $e^{-i\underline{p}\cdot\underline{x}}\underline{m}_-(\hat{p})$ correspond à une hélicité $-1/2$. Quant à la solution $e^{i\underline{p}\cdot\underline{x}}\underline{m}_-(\hat{p})$, son évolution en temps est e^{iEt} et correspond à une rotation dans le sens direct ; il est naturel qu'elle soit associée à l'antiparticule d'hélicité $1/2$ quand le champ est quantifié.

REMARQUES. 1) Les champs $\underline{\psi}$ et $\underline{\widetilde{\psi}}$ sont aussi notés $\underline{\psi}_L$ et $\underline{\psi}_R$ (L pour left et R pour right). 2) L'étude plus détaillée de la quantification de ces champs montre que les opérateurs $a_{\overrightarrow{p}}$, $a^\dagger_{\overrightarrow{p}}$, $b_{\overrightarrow{p}}$, $b^\dagger_{\overrightarrow{p}}$ doivent vérifier des relations d'anticommutation (au lieu de commutation). Cette règle est à l'origine du principe d'exclusion de Pauli relatif aux particules de spin (ou d'hélicité) demi entier (fermions) : impossibilité d'avoir deux particules dans le même état quantique (*cf.* ouvrages spécialisés).

▷ E.D.P. de Dirac ; particules massives de spin $\frac{1}{2}$

L'E.D.P. de Dirac couple des spineurs de types $\underline{\widetilde{\psi}}$ et $\underline{\psi}$, notés ici $\underline{\psi}_L$ et $\underline{\psi}_R$:

$$i\boldsymbol{\partial}\underline{\psi}_L = m\boldsymbol{\psi}_R \quad , \quad i\widetilde{\boldsymbol{\partial}}\underline{\psi}_R = m\boldsymbol{\psi}_L \ .$$

Le lecteur vérifiera comme ci-dessus son invariance et le fait que la conjugaison est une opération interne :

$$\underline{\psi}_D = \begin{pmatrix}\underline{\psi}_R \\ \underline{\psi}_L\end{pmatrix} \text{ solution} \quad \Longrightarrow \quad \underline{\psi}^C_D = \begin{pmatrix}-\boldsymbol{\epsilon}\overline{\underline{\psi}}_L \\ \boldsymbol{\epsilon}\overline{\underline{\psi}}_R\end{pmatrix} \text{ solution} \ , \quad \text{avec } \boldsymbol{\epsilon} = \begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix} \ .$$

L'E.D.P. admet des solutions $\exp \mp i\underline{p} \cdot \underline{x}$ si :

$$\pm(E\mathbf{I} + \overrightarrow{\boldsymbol{\sigma}} \cdot \overrightarrow{p})\underline{\psi}_L = m\underline{\psi}_R \quad , \quad \pm(E\mathbf{I} - \overrightarrow{\boldsymbol{\sigma}} \cdot \overrightarrow{p})\underline{\psi}_R = m\underline{\psi}_L \ .$$

Ceci implique en particulier $(E^2 - \overrightarrow{p}^2)\underline{\psi}_{R,L} = m^2\underline{\psi}_{R,L}$ et donc $E^2 = \overrightarrow{p}^2 + m^2$. Si $\overrightarrow{p} = 0$ ($E = m$, particules au repos), on a $\underline{\psi}_R = \pm\underline{\psi}_L$, et la solution a la forme générale :

$$\underline{\psi}_D = e^{-imt}\left[\alpha_+ \begin{pmatrix} \underline{m}_+ \\ \underline{m}_+ \end{pmatrix} + \alpha_- \begin{pmatrix} \underline{m}_- \\ \underline{m}_- \end{pmatrix}\right] + e^{imt}\left[\overline{\beta}_+ \begin{pmatrix} -\boldsymbol{\epsilon}\overline{\underline{m}}_+ \\ \boldsymbol{\epsilon}\overline{\underline{m}}_+ \end{pmatrix} + \overline{\beta}_- \begin{pmatrix} -\boldsymbol{\epsilon}\overline{\underline{m}}_- \\ \boldsymbol{\epsilon}\overline{\underline{m}}_- \end{pmatrix}\right] \ .$$

L'E.D.P. de Dirac décrit donc, lorsqu'elle est quantifiée, des particules et antiparticules de masse m et de spin 1/2 (avec leurs deux états de spin) par exemple des **électrons** et des **positrons** (*cf.* ouvrages spécialisés).

REMARQUE. Il existe des équations invariantes plus simples, les **E.D.P. de Weyl-Majorana**

$$\boldsymbol{\partial}\underline{\varphi}_L = me^{i\alpha}\boldsymbol{\epsilon}\overline{\underline{\varphi}}_L \quad \text{ou} \quad \widetilde{\boldsymbol{\partial}}\underline{\varphi}_R = me^{i\beta}\boldsymbol{\epsilon}\overline{\underline{\varphi}}_R$$

(phases α, β arbitraires) conduisant à $E^2 = \overrightarrow{p}^2 + m^2$. Mais, comme elles couplent un champ spinoriel à son conjugué, les solutions onde plane sont de la forme $\underline{\varphi}_- e^{-i\underline{p}\cdot\underline{x}} + \underline{\varphi}_+ e^{i\underline{p}\cdot\underline{x}}$ avec à la fois $\underline{\varphi}_-$ et $\underline{\varphi}_+$ non nuls ; elles n'ont pour l'instant pas d'interprétation physique simple. Le lecteur peut vérifier à titre d'exercice que l'E.D.P de Dirac qui concerne un spineur $\underline{\psi}_D$ à quatre composantes se sépare en deux E.D.P. de Weyl-Majorana pour les combinaisons à deux composantes $\underline{\psi}_R \pm \boldsymbol{\epsilon}\overline{\underline{\psi}}_L$.

▷ Masse et fréquence

En physique quantique, la conception "classique" de la masse comme résistance à la modification de la vitesse n'a pas de sens. Une particule libre n'est pas décrite par sa position et sa vitesse mais par sa fonction d'onde $\exp -i(Et - \overrightarrow{p} \cdot \overrightarrow{r})$ dans laquelle E et $\overrightarrow{p}$ figurent comme des fréquences (à 2π près). En physique quantique relativiste, la masse (énergie au repos) est donc une fréquence temporelle. Pour l'interpréter simplement, considérons un espace à une dimension et les équations couplées

$$(\partial_t + \partial_z)\varphi_1 = m\varphi_2 \quad , \quad (\partial_t - \partial_z)\varphi_2 = -m\varphi_1 \ .$$

Pour des solutions $\exp -i(Et - pz)$ elles conduisent à la relation de dispersion $E^2 = p^2 + m^2$ et à $(E - p)\varphi_1 = im\varphi_2$. Pour $m = 0$, φ_1 et φ_2 sont associées aux états de mouvement $c = 1$ et $c = -1$. La masse peut donc être considérée comme une fréquence de transition entre ces états, en analogie avec le couplage des deux états de repos $z = \pm z_0$ de la molécule NH_3 à la section 4.4.2. Si φ_1 et φ_2 sont prises comme amplitudes de probabilité de ces états, on vérifie que la vitesse moyenne définie par $v = (|\varphi_1|^2 - |\varphi_2|^2)/(|\varphi_1|^2 + |\varphi_2|^2)$ est bien la vitesse p/E de la particule de masse m ou la vitesse de groupe dE/dp de l'onde. Une image "classique" de ces transitions est donnée à la figure 20b (avec $\Delta t = m^{-1}$). Le lecteur vérifiera que pour des solutions ne dépendant que de t et z, les E.D.P. de Klein-Gordon et de Dirac peuvent se ramener aux équations pour φ_1 et φ_2, de même que les E.D.P. de Weyl-Majorana en changeant φ_2 en $\overline{\varphi}_2$. Cette interprétation suggère que les états de mouvements (invariants) à la vitesse c sont en physique plus fondamentaux que les états inertiels galiléen de repos ou de vitesse constante $v < c$.

Chapitre 5

Fonctions d'une variable ; analyse des signaux

L'étude des fonctions, longtemps limitée à celles connues pour leur "utilité" (polynômes, logarithme, fonctions trigonométriques...), a connu une première approche générale avec l'introduction vers 1670 du "**calcul infinitésimal**" par Leibnitz et Newton et ses applications à la mécanique et à la géométrie analytique. Ce savoir-faire basé sur les calculs de dérivées, d'intégrales et de développements limités reste très utile en physique.

Cette étude a changé d'objectif quand il s'est agit, dans la seconde moitié du $XIX^{\text{ème}}$ siècle, de distinguer des propriétés "fines" comme celles de continuité, de dérivabilité et d'intégrabilité ou d'établir l'existence (auparavant admise) de fonctions écrites sous forme de séries de Taylor ou de Fourier. Ce développement de l'analyse contemporain de la construction des réels a conduit, les fonctions étant des objets beaucoup plus "riches" que les nombres, à la définition de critères de convergence et d'espaces fonctionnels variés, adaptés à tel ou tel théorème. Au $XX^{\text{ème}}$ siècle un point de vue "intégral" (**"théorie de la mesure"**), initié par Lebesgue (1902), a conduit à abandonner la notion de fonction "définie point par point" pour celle plus générale de distribution (1940-60). Ce point de vue "moderne" n'est pas éloigné de celui de la physique.

En effet, dans la deuxième moitié du $XX^{\text{ème}}$ siècle, s'est dégagée en physique une approche "**traitement du signal**" des fonctions. Elle correspond au développement de méthodes expérimentales générales d'analyse des signaux basées souvent sur des techniques d'intégration : filtrage, détection synchrone (analyse de Fourier), corrélation, échantillonnage, analyse en ondelettes, *etc.* L'intérêt de ces méthodes est d'être applicables à des signaux extrêmement divers : de durée "finie" (décharge d'un condensateur) ou "infinie" (champ électrique dû à une source lumineuse), périodiques (signaux d'un générateur BF) ou chaotiques (bruit de fond d'une résistance), déterministes (régis par une équation différentielle) ou non, d'apparence "continue" ou "discrète" (suite d'impulsions délivrées par un photomultiplicateur), *etc.* Comme on le verra aussi dans ce chapitre, elles s'étendent aux fonctions de plusieurs variables (ondes ou images en optique). De plus en plus elles concernent des signaux numérisés (après échantillonnage et quantification).

Cette approche a contribué a donner en physique une importance primordiale aux **opérations effectuées sur les fonctions** et leurs propriétés mathématiques formelles plutôt qu'à l'appartenance des fonctions à tel ou tel espace fonctionnel. Paradoxalement l'hypothèse la plus raisonnable est de considérer que les fonctions en physique sont a priori les plus "braves" possibles (pourquoi pas indéfiniment dérivables à support compact ?). Les fonctions "singulières" (discontinues, fonction $\delta\ldots$) ou dont le support va à l'infini, très utiles conceptuellement, correspondent à des situations idéalisées (transition "brutale" entre deux milieux, source ponctuelle, ou milieu infini, signal permanent…) ; ces fonctions peuvent toujours être approchées par des "braves" fonctions, non au sens naïf de la convergence point par point mais à celui des distributions (*cf.* exemple de l'impulsion de Dirac).

5.1 SAVOIR-FAIRE CONCERNANT LES FONCTIONS

5.1.1 Graphe et informations sur une fonction

L'intérêt d'un graphe par rapport à une tabulation est de mettre en valeur certaines informations qualitatives importantes.

▷ Domaine et comportement aux bornes

Pour construire le graphe, les premiers points à examiner sont le domaine de la variable x pour lequel la correspondance $x \to y = f(x)$ est bien (univoquement) définie, et le comportement aux bornes de $f(x)$. Par exemple pour $x \to \infty$ on peut chercher si il existe une droite **asymptote** (cas où $f(x)-(ax+b) \to 0$), ou une **branche parabolique** d'axe Ox ou Oy (cas où $\dfrac{f(x)}{x}$ tend vers zéro ou l'infini), par exemple $f(x) = x^\alpha$ pour $0 < \alpha < 1$ ou pour $\alpha > 1$ (figure 1).

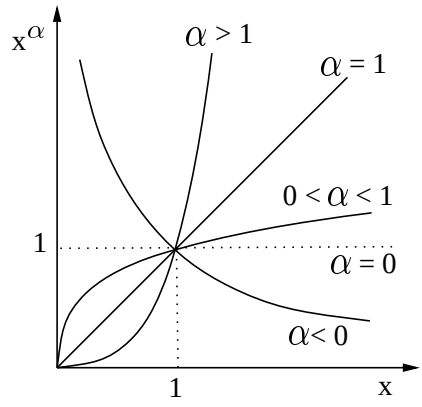


FIGURE 1

Il peut être plus judicieux d'utiliser un **graphe log-log** (c.a.d. $\log|f(x)|$ versus $\log x$) pour faire apparaître une asymptote si $f(x)$ a un comportement en x^α (exemple des diagrammes de Bode pour les fonctions de transfert en électronique) ou un graphe semi-log

(en général $\log|f(x)|$ versus x) pour distinguer par exemple des comportements $f(x) \to 0$ très différents pour $x \to \infty$ tels que x^α ($\alpha < 0$), e^{-x} ou e^{-x^2}. Si x_0 est un **point singulier** ou un **point d'arrêt** du domaine on étudie, lorsque $x \to x_0$ par valeur supérieure ou (et) inférieure, si $f(x)$ devient infinie, par exemple comme $|x - x_0|^\alpha$ pour $\alpha < 0$ ou $\ln|x - x_0|$, ou si $f(x)$ tend vers une constante, par exemple comme $|x - x_0|^\alpha$ pour $\alpha > 0$; dans ce dernier cas il faut préciser la direction de la tangente en x_0 (verticale si $\alpha < 1$ et horizontale si $\alpha > 1$). Souvent cette étude "aux bornes" et quelques valeurs particulières suffisent pour avoir une allure du graphe.

▷ Dérivées

Un complément d'information, là où $f(x)$ a un comportement régulier, est donné par l'étude des dérivées première

$$f'(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}$$

égale, en coordonnées cartésiennes, à la **pente** $\operatorname{tg} \alpha(x)$ de la tangente au graphe, et seconde $f''(x)$ qui permet, grâce au développement de Taylor (*cf.* section 1.1.2)

$$f(x+a) = f(x) + a\, f'(x) + \frac{a^2}{2!}\, f''(x) + \cdots ,$$

de situer le graphe par rapport à la tangente : l'écart en $x + a$ avec la tangente en x est $f(x+a) - \big(f(x) + a\, f'(x)\big) \simeq \dfrac{a^2}{2}\, f''(x)$ pour a petit et $f''(x) \neq 0$. Rappelons que localement la fonction f est croissante si $f'(x) > 0$, décroissante si $f'(x) < 0$ et qu'elle est **stationnaire** (ou extrémale) en x_0 si $f'(x_0) = 0$ (maximum si $f''(x_0) < 0$, minimum si $f''(x_0) > 0$). f est une fonction **concave** si $f''(x) \leq 0$, convexe si $f''(x) \geq 0$ (graphe respectivement en dessous et au dessus de la tangente), et présente un **point d'inflexion** en x_0 si $f''(x)$ s'annule en ce point en changeant de signe.

EXEMPLE. L'**équation d'état de Van der Waals** en variables réduites s'écrit :

$$\Big(p + \frac{3}{v^2}\Big)(3v - 1) - 8t = 0 = F(p, v, t) \ .$$

($p = \dfrac{P}{P_c}$, $v = \dfrac{V}{V_c}$ et $t = \dfrac{T}{T_c}$ sont les valeurs de la pression, du volume et de la température en prenant les coordonnées du point critique comme unités). Les graphes des fonctions $p_t(v) = \dfrac{8t}{3v-1} - \dfrac{3}{v^2}$, pour différentes valeurs fixées de t (isothermes du fluide en coordonnées de **Clapeyron**) (figure 2a) se déduisent des remarques suivantes : $p \to 0^+$ pour $v \to \infty$ (asymptote horizontale $p = 0$) ; $p \to \infty$ pour $v \to \dfrac{1}{3}^+$ (asymptote verticale $v = \dfrac{1}{3}$) ; $p_{t_2}(v) > p_{t_1}(v)$ si $t_2 > t_1$; annulation avec changement de signe de $\dfrac{\mathrm{d}p_t}{\mathrm{d}v}$ sur la courbe en cloche $p = \dfrac{3}{v^2} - \dfrac{2}{v^3}$, par le sommet C ($p = 1, v = 1$) de laquelle passe l'isotherme critique $t = 1$. Pour les autres informations présentes sur le graphe (palier de liquéfaction AB, zones de retard AA' à la vaporisation et BB' à la liquéfaction), le lecteur se reportera à la section 7.4.1 et aux ouvrages de physique.

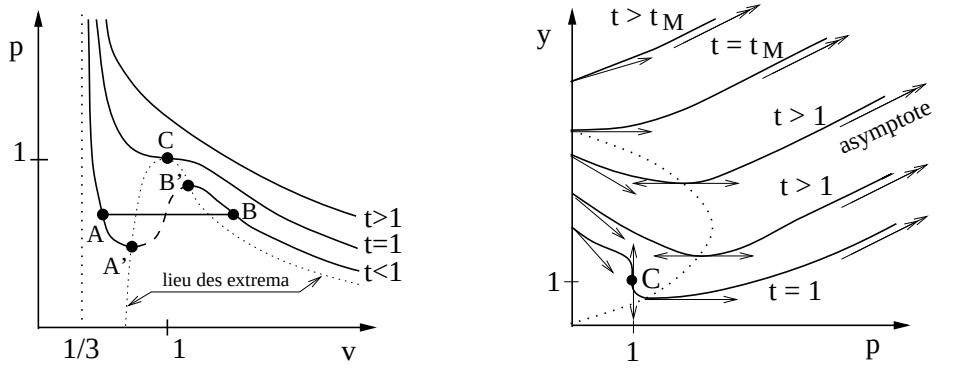


FIGURE 2

En coordonnées d'**Amagat** $(p, y = pv)$ l'équation s'écrit :

$$3y - p + \tfrac{9p}{y} - \tfrac{3p^2}{y^2} - 8t = 0 = G(y, p, t) \ .$$

On en déduit facilement que les isothermes se comportent pour p petit (c.a.d près de $y = \frac{8}{3}t$) comme $y_t \sim \frac{8}{3}t + (\frac{1}{3} - \frac{9}{8t})\, p$, pour p grand (c.a.d. pour $\frac{p}{y} = v^{-1} \sim 3$) comme $y_t \sim \frac{1}{3}p + \frac{8}{3}t$ (asymptotes obliques parallèles) et que leurs pentes s'annulent sur la parabole de Mariotte $y^2 - 9y + 6p = 0$ (obtenue en annulant $\frac{\partial G}{\partial p}$) ; noter aussi que $\frac{\mathrm{d}y_t}{\mathrm{d}p} = v + p_t \frac{\mathrm{d}v}{\mathrm{d}p_t}$ devient infini au point critique C (figure 2b). Pour $t < 1$ les fonctions $y_t(p)$ (comme $v(p_t)$) ne sont plus définies partout de façon univoque.

▷ **Autres informations** (figure 3)

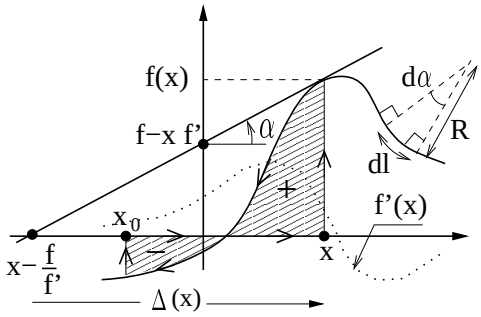


FIGURE 3

Le graphe de $f(x)$ permet aussi de donner une représentation géométrique d'un certain nombre de fonctions reliées à f telles que : les **primitives** $F_{x_0}(x) = \int_{x_0}^{x} f(s)\, ds$, somme des aires algébriques des boucles orientées sur la figure ; la **dérivée logarithmique**

$$\frac{\mathrm{d} \ln |f(x)|}{\mathrm{d}x} = \frac{f'(x)}{f(x)} = \frac{1}{\Delta(x)} \ ;$$

les fonctions $f(x) - xf'(x)$ et $x - \dfrac{f(x)}{f'(x)}$, dont les valeurs sont données par les intersections

de la tangente avec les axes ; la **fonction inverse** f^{-1} définie par la correspondance

$$y = f(x) \longleftrightarrow x = f^{-1}(y)$$

et dont le graphe, ensemble des points de coordonnées (y, x), est le symétrique par rapport à la première bissectrice de celui de f, ensemble des points (x, y) ; on fera attention que f^{-1} n'est définie que si la correspondance entre x et y est biunivoque, par exemple $y = e^x$ et $x = \ln y$, ou $y = \sin x$ et $x = \arcsin y$ pour $x \in]-\pi/2, \pi/2]$).

REMARQUE. Si $y = f(x)$ représente une courbe du plan euclidien, le profil d'une corde vibrante, la section d'un dioptre par un plan..., les quantités intrinsèques à cette courbe, c.a.d. indépendantes du système d'axes orthonormés choisi, sont son **élément de longueur** $\mathrm{d}l = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$ et sa **courbure** $\dfrac{1}{R(x)} = \dfrac{\mathrm{d}\alpha(x)}{\mathrm{d}l}$; $R(x)$ est le rayon du cercle (osculateur) qui approche au mieux la courbe près de x, et $\mathrm{d}\alpha(x)$ l'angle dont tourne la tangente lorsque on se déplace de $\mathrm{d}l$ sur la courbe (*cf.* section 3.2.1). On a

$$\mathrm{d}l = \sqrt{1 + f'^2(x)}\,\mathrm{d}x \quad \text{et} \quad \frac{1}{R(x)} = \frac{f''(x)}{(1 + f'^2(x))^{\frac{3}{2}}} \simeq f''(x) \text{ si } f'^2(x) \ll 1 \ ,$$

car $f''(x) = \dfrac{\mathrm{d}\,\mathrm{tg}\alpha(x)}{\mathrm{d}x} = \big(1 + f'^2(x)\big)\,\dfrac{\mathrm{d}\alpha(x)}{\mathrm{d}x}$; la courbure en $x = 0$ de $y = \dfrac{x^2}{2R}$ est $\dfrac{1}{R}$.

5.1.2 Dérivation, développements limités : principaux résultats

▷ Règles de dérivation (rappels)

Dérivée d'une fonction de fonction : $\dfrac{\mathrm{d}}{\mathrm{d}x} f\left(g\left(x\right)\right) = f'\left(g\left(x\right)\right) g'(x)$ (= 1 si f et g sont inverses l'une de l'autre).

Addition des dérivations "partielles" : lorsque x apparaît à plusieurs endroits dans une expression il faut ajouter les dérivées obtenues en ne faisant varier x qu'à une seule place à la fois par exemple $(uv)' = u'v + uv'$, $(uv)'' = u''v + 2u'v' + uv''$ ou $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_{x+b}^{x+a} f(x,y)\,\mathrm{d}y = f(x, x+a) - f(x, x+b) + \int_{x+b}^{x+a} \frac{\partial f(x,y)}{\partial x}\,\mathrm{d}y$. Si $y(x)$ est défini de manière implicite par $G(y(x), x) = 0$, on obtient y' par $\dfrac{\partial G}{\partial y}\,y' + \dfrac{\partial G}{\partial x} = 0$ (donc $y' = 0$ pour $\dfrac{\partial G}{\partial x} = 0$; *cf.* isothermes de Van der Waals ci-dessus).

Dérivée logarithmique : $\dfrac{\mathrm{d}}{\mathrm{d}x} \ln|f(x)| = \dfrac{f'(x)}{f(x)}$. Elle est très utile pour calculer f' quand f contient des produits et des quotients :

$$\left(f_1^{\alpha_1} f_2^{\alpha_2} \ldots\right)' = \left(f_1^{\alpha_1} f_2^{\alpha_2} \ldots\right)\left(\alpha_1 \frac{f_1'}{f_1} + \alpha_2 \frac{f_2'}{f_2} + \ldots\right) \ ;$$

par exemple le maximum de la **distribution de Maxwell** $P(v) \propto v^2 \exp - \dfrac{1}{2}\dfrac{mv^2}{k_B T}$ est obtenu pour $\dfrac{2}{v} - \dfrac{mv}{k_B T} = 0$ et celui de la **loi de Planck** $u(\nu) \propto \nu^3 \left(\exp\dfrac{h\nu}{k_B T} - 1\right)^{-1}$ pour

$$\frac{3}{\nu} - \frac{h}{k_B T} \exp\frac{h\nu}{k_B T} \left[\exp\frac{h\nu}{k_B T} - 1\right]^{-1} = 0 \text{ soit } h\nu \simeq 3k_B T.$$

EXEMPLES de **dérivées** et de **primitives**. Si on désigne le passage d'une fonction f à sa dérivée f' par $f' \leftharpoondown f$ et à sa primitive la plus simple F par $f \rightharpoonup F$ on a : $e^x \rightleftharpoons e^x$; $\cosh x \rightleftharpoons \sinh x \rightleftharpoons \cosh x$; $\cos x \rightleftharpoons \sin x \rightleftharpoons -\cos x$; $\alpha\, x^{\alpha-1} \rightleftharpoons x^\alpha \rightleftharpoons \dfrac{x^{\alpha+1}}{\alpha+1}$ ($\alpha \neq -1$ et $x > 0$ si α n'est pas entier) ; $x^{-1} \rightleftharpoons \ln|x| \rightleftharpoons x\,\ln|x| - x$; $\dfrac{1}{\cos^2 x} \rightleftharpoons \operatorname{tg} x \rightleftharpoons -\ln|\cos x|$; $\dfrac{1}{\cosh^2 x} \rightleftharpoons \tanh x \rightleftharpoons \ln\cosh x$; $\dfrac{1}{1+x^2} \rightleftharpoons \operatorname{arctg} x = \varphi$ ($x = \operatorname{tg}\varphi$; $\varphi \in \left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$) ; $\dfrac{1}{\sqrt{1-x^2}} \rightleftharpoons \arcsin x = \varphi$ ($x = \sin\varphi$; $\varphi \in \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$) ; $\dfrac{1}{\sqrt{1+x^2}} \rightleftharpoons \operatorname{argsinh} x = \ln\left(x + \sqrt{1+x^2}\right) = \varphi$ ($x = \sinh\varphi$) ; $\dfrac{1}{\sqrt{x^2-1}} \rightleftharpoons \ln\left|x + \sqrt{x^2-1}\right| = \varphi$ ($x = \pm\cosh\varphi$).

EXEMPLES de **développements limités**. Le développement en série de Taylor $f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\,\dfrac{x^n}{n!}$ de quelques fonctions "classiques" s'écrit : $e^x = \sum_{n=0}^{\infty} \dfrac{x^n}{n!}$, $\cosh x = \sum_{n=0}^{\infty} \dfrac{x^{2n}}{(2n)!}$, $\sinh x = \sum_{n=0}^{\infty} \dfrac{x^{2n+1}}{(2n+1)!}$, $\cos x = \sum_{n=0}^{\infty} (-1)^n \dfrac{x^{2n}}{(2n)!}$, $\sin x = \sum_{n=0}^{\infty} (-1)^n \dfrac{x^{2n+1}}{(2n+1)!}$, $(1+x)^\alpha = 1 + \sum_{n=1}^{\infty} \alpha(\alpha-1)\cdots(\alpha-n+1)\,\dfrac{x^n}{n!}$ (attention $|x| < 1$ pour la convergence), et $\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \dfrac{x^n}{n}$ (attention $|x| < 1$).

Remarque : en physique il est rare que la série complète soit utile ; en effet ces développements, qui interviennent lors de calculs de perturbation ne pouvant être conduits que jusqu'à un certain ordre, n'ont d'intérêt que si les premiers termes correctifs sont "petits" (ce qu'il faut toujours vérifier a posteriori) et suffisants (ce qu'on espère mais qu'il est souvent difficile de prouver). En pratique les développements suivants suffisent souvent :

$$e^x = 1+x+\frac{x^2}{2}\ ,\ \cosh x = 1+\frac{x^2}{2}\ ,\ \sinh x = x+\frac{x^3}{3!}\ ,\ \cos x = 1-\frac{x^2}{2}\ ,\ \sin x = x-\frac{x^3}{3!}\ ,$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 \quad \text{et} \quad \ln(1+x) = x - \frac{x^2}{2}\ .$$

Rappelons qu'on peut faire tous les calculs à partir des développements limités à condition de rester cohérent quant aux ordres traités ; par exemple la connaissance de $\cosh x$ et $\sinh x$ à l'ordre 3 permet de déterminer $\tanh x = x\left(1+\dfrac{x^2}{3!}\right)\left(1-\dfrac{x^2}{2}\right) = x - \dfrac{x^3}{3}$ à l'ordre 3, mais $\operatorname{cotanh} x - \dfrac{1}{x} = \dfrac{1}{x}\left(\left(1+\dfrac{x^2}{2}\right)\left(1-\dfrac{x^2}{3!}\right) - 1\right) = \dfrac{x}{3}$ uniquement à l'ordre 1 ; de même $(1-2ax+x^2)^{-\frac{1}{2}} = 1 - \dfrac{1}{2}(-2ax+x^2) + \dfrac{3}{8}(-2ax+x^2)^2 + \cdots = 1 + ax + \dfrac{(3a^2-1)}{2}x^2$ à l'ordre 2 (application section 4.2.5).

5.1.3 Intégration ; cas des fonctions piquées ou rapidement oscillantes

▷ **Règles d'intégration** (rappels)

Changement de variable d'intégration : $\displaystyle\int_{x_1}^{x_2} f(x)\,\mathrm{d}x = \int_{\varphi_1}^{\varphi_2} f\left(x\left(\varphi\right)\right)\,\frac{\mathrm{d}x}{\mathrm{d}\varphi}\,\mathrm{d}\varphi$ à condition de prendre garde que $x(\varphi)$ couvre bien une fois $[x_1, x_2]$ lorsque φ parcourt $[\varphi_1, \varphi_2]$ (exemple : $f(x) = (1+x^2)^{-1}$, $x = \mathrm{tg}\varphi$, $f(x)\,\mathrm{d}x = \mathrm{d}\varphi$, $\varphi \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$).

Intégration par parties : $\displaystyle\int_{x_1}^{x_2} u'v\,\mathrm{d}x = uv\,\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} uv'\,\mathrm{d}x$, formule que l'on peut itérer si on sait intégrer u ; par exemple une primitive de $e^{ax}\,P(x)$ est $e^{ax}\left(\dfrac{P(x)}{a} - \dfrac{P'(x)}{a^2} + \dfrac{P''(x)}{a^3}\cdots\right)$, formule utile si $P(x)$ est un polynôme, et applicable à $e^{iax}P(x)$ pour obtenir la primitive de $\cos ax\,P(x)$.

Dérivation par rapport à des paramètres introduits ou choisis astucieusement (exemples) :

1) $\displaystyle\int_0^\infty e^{-x}x^n\,\mathrm{d}x = \left[-\frac{\mathrm{d}}{\mathrm{d}a}\right]^n \left[\int_0^\infty e^{-ax}\,\mathrm{d}x = \frac{1}{a}\right]\Big|_{a=1} = n!\,;$

2) à l'aide de

$$\int_{-\infty}^{\infty} e^{-ax^2}x^{2n}\,\mathrm{d}x = \left[-\frac{\mathrm{d}}{\mathrm{d}a}\right]^n \left[\int_{-\infty}^{\infty} e^{-ax^2}\,\mathrm{d}x = \sqrt{\frac{\pi}{a}}\right] = \sqrt{\pi}\,\frac{1}{2}\,\frac{3}{2}\cdots\frac{2n-1}{2}\,a^{-\left(n+\frac{1}{2}\right)}\ ,$$

on obtient les moments de la **loi gaussienne** :$< x^2 >= \sigma^2$, $< x^4 >= 3\sigma^4$ et de façon générale

$$< x^{2n} >= (2\pi\sigma^2)^{-\frac{1}{2}} \int_{-\infty}^{\infty} x^{2n}\,e^{-\frac{x^2}{2\sigma^2}}\,dx = 1\times 3\times\cdots\times(2n-1)\,\sigma^{2n}\ ;$$

3) la moyenne du module de la vitesse (loi de Boltzmann) est, avec $a = \dfrac{m}{2k_BT}$:

$$\begin{aligned} < v > &= \left(\int_0^\infty e^{-av^2}v^3\,\mathrm{d}v\right)\left(\int_0^\infty e^{-av^2}v^2\,\mathrm{d}v\right)^{-1} \\ &= \left(-\frac{\mathrm{d}}{\mathrm{d}a}\int_0^\infty e^{-av^2}v\,\mathrm{d}v\right)\left(-\frac{\mathrm{d}}{\mathrm{d}a}\int_0^\infty e^{-av^2}\,\mathrm{d}v\right)^{-1} \\ &= \left(-\frac{\mathrm{d}}{\mathrm{d}a}\,\frac{1}{2a}\right)\left(-\frac{\mathrm{d}}{\mathrm{d}a}\,\frac{1}{2}\sqrt{\frac{\pi}{a}}\right)^{-1} = \frac{2}{\sqrt{\pi}}\,a^{-\frac{1}{2}}\ . \end{aligned}$$

Le lecteur trouvera de nombreux autres exemples en physique statistique.

▷ Approximation de la "méthode du col" (figure 4a)

Si $f(x)$ est maximale en $x^* \in]x_1, x_2[$ on a pour K grand

$$I = \int_{x_1}^{x_2} g(x)\, e^{K\, f(x)}\, \mathrm{d}x \simeq g(x^*) \sqrt{\frac{2\pi}{K|f''(x^*)|}}\, e^{K\, f(x^*)} \quad (K \gg 1)\ ,$$

ou en pratique $\ln I \simeq K \max f(x)$. Ce résultat se justifie en remarquant que $\exp\bigl(K f(x)\bigr)$ est très piquée autour de x^* et qu'on peut par conséquent remplacer $g(x)$ par $g(x^*)$, $f(x)$ par $f(x^*) + \dfrac{(x-x^*)^2}{2}\, f''(x^*)$ et considérer que l'intégrale de la gaussienne va de $-\infty$ à $+\infty$.

EXEMPLE : $n! = \displaystyle\int_0^\infty x^n\, e^{-x}\, \mathrm{d}x \overset{x \to ny}{=} n \int_0^\infty e^{n(\ln n + \ln y - y)}\, \mathrm{d}y \simeq \sqrt{2\pi n}\, n^n\, e^{-n}$, d'où la **formule de Stirling**

$$\ln n! \simeq n(\ln n - 1)$$

pour n grand (qu'on a aussi par $\ln n! = \ln 1 + \cdots + \ln n \simeq \int_1^n \ln x\, \mathrm{d}x$).

REMARQUES. 1) On a de même $\int e^{\frac{1}{k_B}\left(S_1(U_1)+S_2(U-U_1)\right)} \mathrm{d}U_1 \simeq \exp\Bigl(\frac{1}{k_B} \max_{U_1}\bigl[S_1(U_1) + S_2(U - U_1)\bigr]\Bigr)$: le nombre total d'états microscopiques pour un système de deux corps dont l'énergie totale U est fixée est donc pratiquement égal, pour le **calcul de l'entropie**, au nombre d'états microscopiques correspondant à l'état macroscopique d'équilibre (*cf.* section 1.1.4). 2) Si f a plusieurs maxima seul le plus grand "compte". 3) La méthode du col et l'expression de $\ln I$ s'étendent aux fonctions de plusieurs variables.

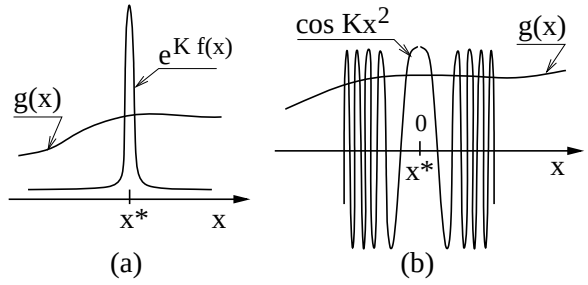


FIGURE 4

▷ Approximation de la phase stationnaire

La formule précédente s'applique au cas de fonctions rapidement oscillantes (remplacement de K par iK). La région de x autour de x^* qui contribue à I est alors celle où $\exp\bigl(iK\, f(x)\bigr)$ oscille le moins (figure 4b). Attention : $Kf''(x^*)$ peut ici être positif ou négatif et $\displaystyle\int_{-\infty}^{\infty} e^{\pm i|Kf''(x^*)|\frac{(x-x^*)^2}{2}} \mathrm{d}x = \sqrt{\frac{2\pi}{|Kf''(x^*)|}}\, e^{\pm i\frac{\pi}{4}}$.

EXEMPLE 1. Toute combinaison d'ondes planes $|A(k)|e^{-i\left(\omega(k)t - kx + \varphi(k)\right)}$ telles que $|A(k)|$ varie lentement mais dont chaque terme de la phase varie "vite" (donc en particulier pour t et x grands) s'écrit, à un facteur près,

$$\int_{\mathcal{D}} |A(k)|e^{-i\left(\omega(k)t - kx + \varphi(k)\right)}\, \mathrm{d}k \propto |A(k^*)|e^{-i\left(\omega(k^*)t - k^* x + \varphi(k^*)\right)}\ ,$$

où $k^*(x,t) \in \mathcal{D}$ (vecteur d'onde local) est tel que $v_g(k^*)t - x + \dfrac{\mathrm{d}\varphi}{\mathrm{d}k}(k^*) = 0$ (**séparation des ondes selon leur vitesse de groupe** $v_g(k^*) = \dfrac{\mathrm{d}\omega}{\mathrm{d}k}(k^*)$). Si il n'existe pas de valeur de k^* satisfaisant cette condition l'onde résultante (l'intégrale) est "négligeable". Dans le cas particulier où $\mathcal{D}$ est un domaine étroit autour de k_0 (et $k_0 = k^*$ sinon il n'y a pas d'onde résultante), on a affaire à un **paquet d'ondes** localisé à l'instant t autour du point $x = v_g(k_0)t + \dfrac{\mathrm{d}\varphi}{\mathrm{d}k}(k_0)$; il se déplace à la vitesse de groupe $v_g(k_0)$ (*cf.* une autre approche à la section 8.2.1).

EXEMPLE 2. Dans l'intégrale d'Huygens-Fresnel $A(M) = \iint A(m)\dfrac{e^{i\frac{2\pi}{\lambda}mM}}{i\lambda mM}\,\mathrm{d}^2m$ étendue à une surface d'onde (argument de $A(m)$ fixé) le point m^* qui contribue essentiellement à $A(M)$ dans l'approximation de l'**optique géométrique** (λ "petit") est celui pour lequel m^*M est stationnaire, c'est-à-dire tel que $\overrightarrow{m^*M}$ est orthogonal à la surface d'onde : m^*M est alors un rayon. Si il y a plusieurs points m^* (cas où M est proche d'une caustique ; *cf.* section 7.3.2), il faut ajouter leurs contributions (qui donnent alors des franges d'interférences).

5.1.4 Concavité de l'entropie ; travail maximum ; transitions de phase

On considère le cas simple où l'entropie S ne dépend que de l'énergie interne U (et pas du volume V : systèmes quasi incompressibles tels que des solides, des liquides...). Le but est d'illustrer sur des graphes le lien entre le second principe et les propriétés (admises ici) de croissance monotone $\left(\dfrac{\mathrm{d}S}{\mathrm{d}U} = \dfrac{1}{T} > 0\right)$ et de concavité $\left(\dfrac{\mathrm{d}^2S}{\mathrm{d}U^2} = -\dfrac{1}{T^2}\dfrac{\mathrm{d}T}{\mathrm{d}U} \leq 0\right)$ de S, ainsi que quelques résultats classiques (travail maximum extrait, transitions de phase...) conséquences directes de ces propriétés. Dans la suite on utilise les grandeurs massiques $u = \dfrac{U}{m}$ et $s = \dfrac{S}{m}$. (Remarque : pour des systèmes compressibles et des transformations à pression constante, par exemple la vaporisation d'un corps pur, il faudrait considérer l'entropie comme fonction de la variable enthalpie H, qui est la seule variable dont dépend alors S, et utiliser $S'(H) = T^{-1}$.)

▷ Thermalisation et travail maximum

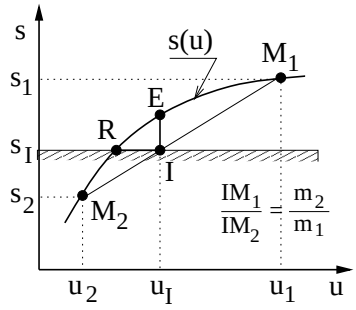


FIGURE 5

EXEMPLE. Sur la figure 5 les points M_i $(i = 1, 2)$ de coordonnées (u_i, s_i) correspondent aux états d'équilibre initiaux de deux corps de masse m_i constitués d'un même

matériau, l'un ($i = 1$) chaud et l'autre ($i = 2$) froid. Ils se trouvent sur le graphe de la fonction $s(u)$ caractéristique du matériau. Les énergie et entropie massiques moyennes du système des deux corps sont $u_I = \dfrac{m_1u_1 + m_2u_2}{m_1 + m_2}$ et $s_I = \dfrac{m_1s_1 + m_2s_2}{m_1 + m_2}$, coordonnées du point I représentatif de l'état initial hors équilibre du système.
Si le système est isolé, la mise en contact thermique des corps conduit chacun d'eux dans le même état d'équilibre E sur $s(u)$ tel que $u_E = u_I$ (énergie conservée, premier principe). L'augmentation d'entropie ($s_E > s_I$: second principe) est donc associée à la concavité de $s(u)$.
Si le système est en contact avec un système "mécanique" ($S_{\text{meca}} = 0$; *cf.* section 1.1.4), l'état d'équilibre final F est sur la courbe $s(u)$ à l'abscisse $u_F = u_I + \dfrac{W}{m_1 + m_2}$ (premier principe ; W travail reçu) avec $s_F \geq s_I$ (second principe). On voit sur la figure que le travail maximum que le système peut fournir est $(-W)_{max} = (m_1+m_2)(u_I-u_R)$, R étant le point de $s(u)$ tel que $s_R = s_I$ (évolution réversible). Si $u = u_0 + c(T - T_0)$, alors $\dfrac{\mathrm{d}u}{\mathrm{d}s} = T$ implique $s = s_0 + c \ln \dfrac{T}{T_0}$, et pour $m_1 = m_2 = m$ on obtient aussitôt $T_E = \dfrac{T_1 + T_2}{2}$, $T_R = \sqrt{T_1T_2}$ et $(-W)_{max} = 2mc(T_E - T_R)$.

▷ Transition de phase

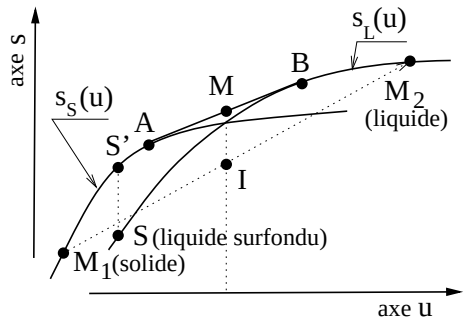


FIGURE 6

EXEMPLE 1. La figure 6 montre les graphes (concaves et croissants) des entropies $s_{\scriptscriptstyle S}(u)$ et $s_{\scriptscriptstyle L}(u)$ des **phases solide et liquide** d'un même corps, en supposant que $s_{\scriptscriptstyle L}(u) > s_{\scriptscriptstyle S}(u)$ pour u grand (état liquide stable à haute température) et $s_{\scriptscriptstyle S}(u) > s_{\scriptscriptstyle L}(u)$ pour u petit (état solide stable à basse température). La réunion des parties stables des deux graphes n'étant pas concave, elle ne peut représenter la fonction $s(u)$ du corps. Soit AB leur tangente commune de pente T_0^{-1} où T_0 est la température des phases solide (en A) et liquide (en B). Si on part du point I représentatif d'un mélange (hors équilibre) de solide dans l'état M_1 et de liquide dans l'état M_2 ($\dfrac{M_1I}{M_2I} = \dfrac{m_{\scriptscriptstyle L}}{m_{\scriptscriptstyle S}}$) tel que

$u_A < u_I < u_B$, l'évolution spontanée vers l'équilibre se fait par transformation d'une phase dans l'autre : M_1 se déplace donc sur $s_{_S}(u)$ et M_2 sur $s_{_L}(u)$ en respectant pour I (mélange) les conditions $u =$ constante et s maximum à l'équilibre ; ceci amène le point I en M sur AB. La courbe représentative de $s(u)$ (arc M_1A, segment AB, arc BM_2) est appelée **enveloppe concave** des deux graphes. Le segment AB correspond au domaine de coexistence des deux phases (palier de **fusion** à T_0), et $L = u_B - u_A = T_0(s_B - s_A)$ est la chaleur latente massique de fusion.
Remarques. 1) **Métastabilité**. Expérimentalement on observe le graphe de $s_{_L}(u)$ au delà de B (existence de liquide surfondu tel que S). Si on introduit un germe, une évolution irréversible amène S en S' ($u_S = u_{S'}$, $s_{S'} > s_S$) ; l'état final est soit du solide (cas de la figure), soit un mélange liquide-solide. 2) Des graphes de $s_{_S}(u)$ et de $s_{_L}(u)$, on déduit les **énergies libres** $f_{S,L}(T) = u - Ts_{_{S,L}}(u)$, intersections de la tangente avec l'axe des u (*cf.* section 5.1.1) ; on voit sur la figure qu'à T fixé ($s'(u)$ fixé) la phase stable est celle qui a la plus petite énergie libre.

EXEMPLE 2. **Paramagnétisme et ferromagnétisme**. Pour un système de N moments magnétiques alignés selon Oz et pouvant prendre les valeurs $\pm\mu$, le nombre d'états microscopiques correspondant à un **taux d'aimantation** $x = \frac{N_+ - N_-}{N} = \frac{M}{N\mu}$ est $\Omega = \frac{N!}{N_+!\,N_-!}$. En utilisant la formule de Stirling $\ln N! = N\,\ln N - N$, on a :

$$S(x) = k_B\,\ln\Omega = -Nk_B\,\left(\tfrac{1+x}{2}\,\ln\tfrac{1+x}{2} + \tfrac{1-x}{2}\,\ln\tfrac{1-x}{2}\right)\ .$$

Le graphe de $S(x)$ est représenté sur la figure 7 : $\frac{\mathrm{d}S}{\mathrm{d}x} = -\frac{Nk_B}{2}\,\ln\frac{1+x}{1-x}$ est infini en $x = \pm1$ et se comporte comme Nk_Bx au voisinage de 0. Soit

$$U(x) = -N\,\left(\mu Bx + \tfrac{1}{2}\,Kx^2\right)$$

l'énergie du système (figure 8) ; le terme $-N\mu Bx$ rend exactement compte des energies de couplage $\mp\mu B$ de chaque moment $\pm\mu$ avec un champ magnétique extérieur B selon Oz ; le second terme rend compte des énergies de couplage de chaque moment avec ses voisins, mais de façon approchée, les moments individuels étant remplacés par des moments moyens (théorie de champ moyen ; *cf.* section 10.2.3). $K > 0$ correspond aux corps ferromagnétiques et $K = 0$ aux corps paramagnétiques.

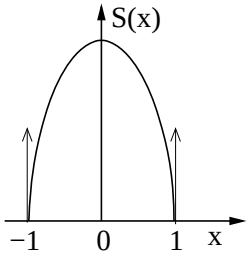


FIGURE 7

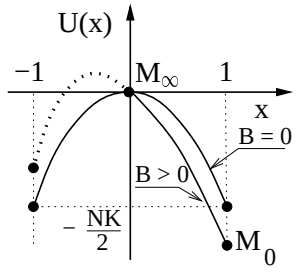


FIGURE 8

L'observation des graphes de $S(x)$ et de $U(x)$ conduit aux graphes de $S(U)$ (figures 9a pour le paramagnétisme et 9b pour le ferromagnétisme). Dans le cas ferromagnétique et pour $B = 0$, on vérifie que lorsque U passe de la valeur $-\frac{NK}{2}$ à 0, la température $T = \frac{\mathrm{d}U}{\mathrm{d}x}\,\left(\frac{\mathrm{d}S}{\mathrm{d}x}\right)^{-1}$ passe de 0 à $T_c = \frac{K}{k_B}$ et $|x|$ passe de 1 à 0 : il existe donc une aimantation spontanée pour $T < T_c$; pour $T > T_c$ il faut considérer que $x = 0$ (*cf.* la limite $B \to 0$ ci-dessous). Pour $B > 0$, la partie physiquement acceptable du graphe

de $S(U)$ correspond au passage de l'énergie U de $-N\mu B - N\frac{K}{2}$ (énergie la plus basse, $x = 1$, $T = 0$, point M_0) à 0 ($x = 0$, T infini, point M_∞) ; la partie en pointillé ne l'est pas car pour $U < 0$ elle ne maximise pas S, et pour $U > 0$ elle ne vérifie pas à la fois $\frac{\mathrm{d}S}{\mathrm{d}U} > 0$ et $\frac{\mathrm{d}^2S}{\mathrm{d}U^2} < 0$.
Remarque : on retrouve ces résultats en représentant les graphes de $F_T(x) = U(x) - TS(x)$; à T fixée la phase stable est celle qui minimise F_T (figure 10 pour $B = 0$).

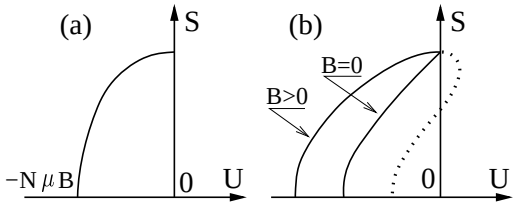


FIGURE 9

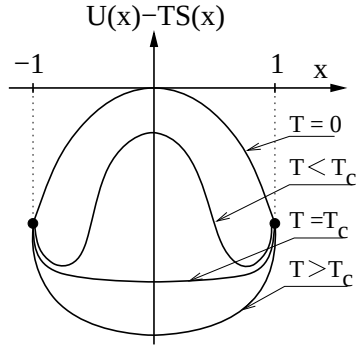


FIGURE 10

Un calcul simple montre que l'égalité $\frac{\mathrm{d}U}{\mathrm{d}x} = T\,\frac{\mathrm{d}S}{\mathrm{d}x}$ donne l'**équation d'état** :

$$x = \tanh\left(\frac{\mu B}{k_B T} + x\frac{T_c}{T}\right) \qquad (k_B T_c = K)\ .$$

La courbe d'aimantation $x(T, B = 0)$ (figure 11a) et les isothermes $x_T(B)$ (figure 11b) s'obtiennent en étudiant les intersections de la droite $y = x$ respectivement avec les courbes $y = \tanh x\frac{T_c}{T}$, dont la pente à l'origine vaut $\frac{T_c}{T}$ (figure 12a), et avec les courbes $y = \tanh\left(\frac{\mu B}{k_B T} + x\frac{T_c}{T}\right)$, qui se déduisent les unes des autres par translation lorsque B varie (figure 12b). Un développement limité près de $x = 0$ et $B = 0$ donne pour l'équation d'état $x\left(1 - \frac{T_c}{T}\right) = \frac{\mu B}{k_B T} - \frac{1}{3}\left(\frac{\mu B}{k_B T} + x\frac{T_c}{T}\right)^3 + \cdots$. Il permet d'obtenir facilement les comportements critiques près de $T = T_c$: $x(T, B = 0) = \sqrt{3\frac{T_c - T}{T_c}}$ ($T < T_c$), $|x_{T_c}(B)| \propto B^{\frac{1}{3}}$ (isotherme critique), $\left.\frac{\mathrm{d}x_T}{\mathrm{d}B}\right|_{B=0} = \frac{\mu}{k_B(T - T_c)}$ pour $T > T_c$, et $\frac{\mu}{2k_B(T_c - T)}$ pour $T < T_c$ (température de Curie).

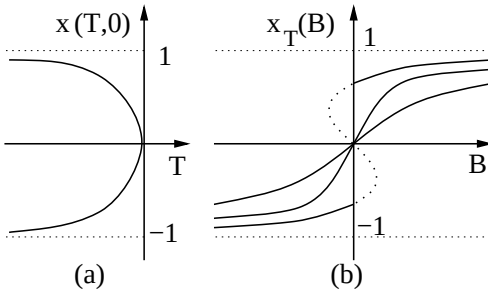


FIGURE 11

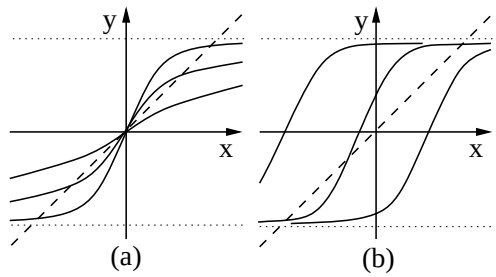


FIGURE 12

5.2 OPÉRATIONS SUR LES FONCTIONS ; ANALYSE DE DIRAC

5.2.1 Opérations sur les fonctions

Les principales opérations sur les fonctions $f(x)$, ou signaux $f(t)$, sont les suivantes.

▷ **La dilatation** $\quad f(x) \to [\mathrm{D}_k f](x) = f\Big(\dfrac{x}{k}\Big) \quad k > 0$

Comme $f\Big(\dfrac{x}{k}\Big)$ prend la valeur $f(x_0)$ en $x = kx_0$, cette opération se traduit par une homothétie de rapport k du graphe de f suivant l'axe des x. Le signal $f\Big(\dfrac{t}{k}\Big)$ varie plus lentement que $f(t)$ si $k > 1$ et plus vite si $0 < k < 1$ (*cf.* cos ωt et cos $\dfrac{\omega}{k}t$ ou figure 15). Dans l'opération l'aire sous le graphe est multipliée par k.

▷ **La parité** $\quad f(x) \to [\mathrm{P}f](x) = f(-x) \quad (\mathrm{P}^2 = I)$

Les fonctions paires $f(-x) = f(x)$ et impaires $f(-x) = -f(x)$ sont fonctions propres de P associées aux valeurs propres 1 et -1. Toute fonction est la somme $f = f_+ + f_-$ de telles fonctions $f_\pm(x) = \dfrac{f(x) \pm f(-x)}{2}$. Une conséquence importante est que tout problème physique linéaire possédant la symétrie $x \to -x$ admet une base de solutions ayant chacune une parité bien définie ; en effet si f est solution, Pf l'est aussi par raison de symétrie, et donc f_+ et f_- aussi par linéarité. Dans le cas non dégénéré une seule de ces fonctions est non nulle (*cf.* par exemple les fonctions propres de l'oscillateur quantique, section 4.4.3).

▷ **La translation** $\quad f(x) \to [\mathrm{T}_a f](x) = f(x - a)$

Comme $f(x - a)$ prend la valeur $f(x_0)$ en $x_0 + a$ cette opération se traduit par une translation a du graphe de f selon x. Le signal $f(t - \tau)$ a un **retard** (algébrique) de τ par rapport à $f(t)$ (*cf.* figure 15).

▷ **La périodisation** $\quad f(x) \to f_{\mathrm{per}}(x) = \displaystyle\sum_{n=-\infty}^{\infty} f(x - na)$

REMARQUE : toute **fonction périodique** de période a, c.a.d. pour laquelle T_a est la plus petite translation la laissant invariante, peut s'écrire comme la périodisée $f_{\mathrm{per}}(x)$ d'une fonction $f(x)$. Par exemple pour $f(x)$ on peut prendre toute restriction de la fonction périodique à un intervalle $[x_0, x_0+a[$; mais il existe une infinité d'autres décompositions, car rien n'interdit que les supports des fonctions $f(x - na)$ se recouvrent (penser à un circuit RC soumis à une suite périodique de créneaux et aux réponses à chaque créneau).

▷ **L'échantillonnage** $\quad f(x) \to f_{\mathrm{ech}}(x) = a \displaystyle\sum_{n=-\infty}^{\infty} f(na)\, \delta(x - na)$

Cette opération étudiée à la section 5.3.2. revient à limiter la connaissance de f au sous ensemble dénombrable des points $x_n = na$. On verra que dans cette "discrétisation" l'information sur f peut cependant ne pas être perdue.

▷ **Le produit** avec une fonction g $\quad f(x) \to f(x)\,g(x)$

EXEMPLES : la restriction d'une fonction à un intervalle $\left[x_0 - \frac{a}{2}, x_0 + \frac{a}{2}\right]$ qui correspond à la multiplication par $\Pi\Big(\frac{x - x_0}{a}\Big)$ où $\Pi(x)$ est la **fonction porte** (figure 13), ou encore la multiplication de $f(t)$ par $\cos\omega_0 t$ qui intervient dans l'opération de modulation d'amplitude $f(t) \to a\big(1 + m f(t)\big)\cos\omega_0 t$. A deux dimensions, la **diffraction** par une ouverture de coefficient de transmission $t(x, y)$ dans le plan $z = 0$ correspond à la multiplication de l'onde incidente dans ce plan par $t(x, y)$.

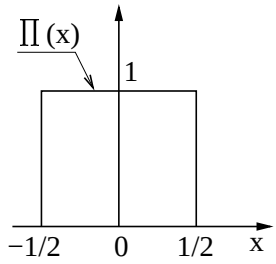


FIGURE 13

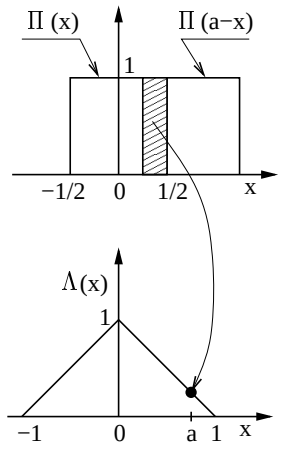


FIGURE 14

▷ **La convolution** avec une fonction g $\quad f(x) \to \big[g * f\big](x) = \displaystyle\int_{-\infty}^{\infty} f(y)\,g(x - y)\,\mathrm{d}y$

Le produit de convolution est commutatif $f * g = g * f$ et associatif $f * (g * h) = (f * g) * h$. Pour le translater ou le dériver il suffit de translater ou dériver un terme du produit. On verra qu'en physique le filtrage se traduit par la convolution (*cf.* section 5.2.4.). Un exemple est la **moyennisation** de f sur un intervalle de largeur a :

$$f(x) \to f_{\text{moyennée}}(x) = \frac{1}{a}\int_{x-a/2}^{x+a/2} f(y)\,\mathrm{d}y = \frac{1}{a}\int_{-\infty}^{+\infty} f(y)\,\Pi\Big(\frac{x - y}{a}\Big)\,\mathrm{d}y\ .$$

De façon générale la convolution avec $\Pi(x)$ a un effet de lissage sur les fonctions et étale leur support. Ainsi alors que Π est discontinue la **fonction "triangle "** $\Lambda = \Pi * \Pi = \Pi^{*2}$ est continue (figure 14), Π^{*3} est partout une fois dérivable... et on verra au chapitre 10 que $\Big(\sqrt{N}\,\Pi\big(\sqrt{N}\,x\big)\Big)^{*N}$ tend vers une gaussienne lorsque $N \to \infty$. En plus de $\Pi * \Pi = \Lambda$ d'autres exemples classiques (en optique, probabilités...) sont $L_a * L_b = L_{a+b}$ ($L_a(x) = \frac{a}{\pi}\,\frac{1}{a^2 + x^2}$ fonction **lorentzienne**) et

$$G_a * G_b = G_{\sqrt{a^2+b^2}} \qquad \Big(G_a(x) = \frac{1}{\sqrt{2\pi a^2}}\,e^{-\frac{x^2}{2a^2}}\ \text{ fonction } \textbf{gaussienne}\Big)\ .$$

REMARQUE. Pour des signaux nuls pour $t < 0$ (*cf.* **électronique**) la convolution s'écrit :

$$\big[g * f\big](t) = \int_0^t f(s)\, g(t-s)\, \mathrm{d}s \ .$$

▷ **La corrélation** avec une fonction g $\quad f(x) \to \Gamma_{f,g}(a) = \displaystyle\int_{-\infty}^{\infty} f(x)\, \overline{g}(x-a)\, \mathrm{d}x = (\mathrm{T}_a g, f)$

La fonction de corrélation $\Gamma_{f,g}$ vérifie :

$$\Gamma_{f,g}(a) = \overline{\Gamma_{g,f}(-a)} \quad \text{et} \quad |\Gamma_{f,g}(a)| \leq \sqrt{\Gamma_{f,f}(0)\, \Gamma_{g,g}(0)} \ \text{(inégalité de Schwarz)} \ .$$

Si $f = g$, $\Gamma_{f,f}(a)$ notée $\Gamma_f\,(a)$ est maximum en 0. Si f et g sont des signaux périodiques ou stationnaires, l'intégrale de $-\infty$ à ∞ (non définie) est remplacée par une moyenne temporelle (*cf.* section 5.3.3.).

REMARQUE : $\Gamma_{f,g}(a)$ est une mesure du "**recouvrement**" (des graphes) des fonctions f et $\mathrm{T}_a g$. Cette notion se retrouve dans différentes méthodes d'analyse des signaux. Par exemple dans l'**analyse de Dirac** (*cf.* section 5.2.3), les valeurs $f(t_0)$ mesurent les recouvrements du signal $f(t)$ avec les "fonctions piquées" $\delta(t-t_0)$ (corrélation de f avec δ), dans l'**analyse de Fourier** les coefficients $\hat{f}(\nu)$ mesurent ceux avec les fonctions $e^{i2\pi\nu t}$ de fréquence ν. Dans une **analyse "temps-fréquence"** on considère les recouvrements avec des fonctions $\varphi(t-t_0)\, e^{i2\pi\nu t}$ (par exemple φ fonction porte) et dans l'**analyse en ondelettes** ("temps-échelle") ceux avec les ondelettes $\varphi(\nu(t-t_0))$ (φ fonction d'intégrale nulle ; figure 15).

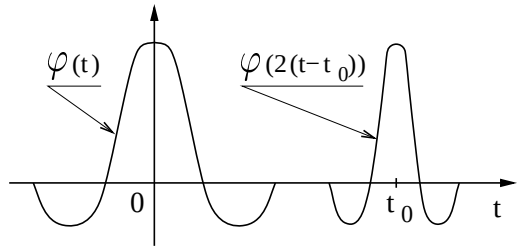


FIGURE 15

▷ Linéarité et stationnarité des opérations

Considérées comme des systèmes entrée-sortie (*cf.* section 2.4.2), toutes ces opérations sont linéaires (sauf $f \to \Gamma_f$ et la modulation d'amplitude). Parmi elles la translation, la périodisation et la convolution sont stationnaires (invariantes par translation). On rencontre aussi en physique de nombreuses opérations non linéaires qui peuvent être stationnaires (détection quadratique $f \to f^2$, "redressement" $f \to |f|$, décalage de zéro $f \to f+$constante, quantification...) ou non (modulations d'amplitude et de fréquence...).

5.2.2 Impulsion de Dirac (« fonction delta ») ; exemples mécaniques

La fonction $\delta(x)$ (ou $\delta(\vec{r}) = \delta(x)\, \delta(y)\, \delta(z)$ à trois dimensions) ou l'impulsion $\delta(t)$ ont une signification physique intuitive. Elles servent à décrire des **grandeurs localisées** sur l'axe x en $x = 0$ (ou dans l'espace : charge électrique ou source lumineuse ponctuelles

situées à l'origine), ou localisées dans le temps (signal bref à $t = 0$). La représentation de toute fonction comme "somme continue" de fonctions delta (analyse de Dirac) est implicite dans de nombreux domaines de la physique.

▷ Définitions

Pratiquement $\delta(x)$ est la limite (au sens indiqué ci-dessous)

$$\delta(x) = \lim_{\epsilon \to 0} \frac{1}{|\epsilon|}\, g\Big(\frac{x}{|\epsilon|}\Big) \quad \text{avec} \quad \int_{-\infty}^{\infty} g(x)\,\mathrm{d}x = 1$$

pour toute fonction $g(x)$ (d'intégrale égale à 1), par exemple $\Pi(x)$ (figure 16), $\Lambda(x)$, $\mathrm{sinc}(\pi x)$, gaussienne, lorentzienne, *etc.* Dans les calculs concrets cette limite intervient toujours dans des intégrales avec des "braves" fonctions $\varphi(x)$, par exemple :

$$\int_{-\infty}^{\infty} \frac{1}{|\epsilon|}\, g\Big(\frac{x-a}{|\epsilon|}\Big)\, \varphi(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} g(y)\, \varphi(a+\epsilon y)\,\mathrm{d}y \quad \to \quad \varphi(a)\ .$$

Tout se passe comme si $\delta(x-a)$ (d'aire égale à 1) avait pour support le point $x = a$.

REMARQUE : pour représenter le graphe de $C\,\delta(x-a)$, on dessine soit une fonction très piquée en $x = a$ dont on précise l'aire C, soit une flèche verticale proportionnelle à C. La fonction "cha" $\text{⨿}\,(x) = \sum_{n=-\infty}^{\infty} \delta(x-n)$ (périodisée de $\delta(x)$ de période 1) s'appelle **"peigne" de Dirac**.

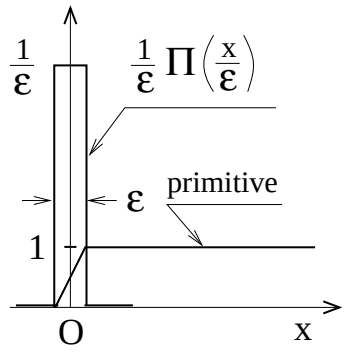


FIGURE 16

▷ Propriétés ; dérivées de fonctions discontinues

On déduit de la définition et de la "localisation" en a de $\delta(x-a)$:

$$\int_{-\infty}^{\infty} \delta(x)\,\mathrm{d}x = 1 \quad ; \quad \delta\Big(\frac{x}{k}\Big) = |k|\delta(x) \quad \text{donc } \delta(x) \text{ paire} \quad ; \quad f(x)\,\delta(x-a) = f(a)\,\delta(x-a)\ .$$

Comme de plus $\displaystyle\int_{-\infty}^{x} \frac{1}{|\epsilon|}\, g\Big(\frac{y}{|\epsilon|}\Big)\,\mathrm{d}y$ tend vers la **fonction de Heaviside (ou échelon)**, $H(x) = 0$ si $x < 0$ et $H(x) = 1$ si $x > 0$ (figure 16), on a aussi :

$$\delta(x) = \frac{\mathrm{d}}{\mathrm{d}x} H(x)\ .$$

Plus généralement, au sens des distributions, on peut dériver une fonction $f(x)$ ayant une discontinuité en a (c.a.d. la limite d'une fonction à variation arbitrairement rapide entre a^- et a^+) et dérivable ailleurs : on ajoute à la dérivée habituelle l'impulsion $\big(f(a^+) - f(a^-)\big)\,\delta(x-a)$ (dont l'intégrale entre a^- et a^+ donne la discontinuité). On notera enfin que $\int_{-\infty}^{\infty}\delta(x)\,\mathrm{d}x = 1$ entraîne que la dimension de $\delta(x)$ est $[x]^{-1}$.

▷ Applications mécaniques ; forces et percussions

Intuitivement un mouvement de vitesse continûment variable $v(t)$ peut être approché par un mouvement dont la vitesse $v_p(t)$ est constante par paliers, avec des sauts Δv_n aux instants t_n (figure 17). La force correspondante ne peut évidemment pas être nulle. En fait $F_p(t) = m\dfrac{\mathrm{d}v_p}{\mathrm{d}t} = \displaystyle\sum_n m\,\Delta v_n\,\delta(t-t_n)$ est une suite d'impulsions mécaniques (**percussions** représentées par des flèches sur la figure).

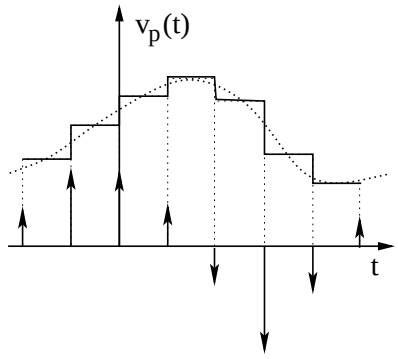


FIGURE 17

Inversement pour résoudre de façon approchée l'équation différentielle $m\dfrac{\mathrm{d}v}{\mathrm{d}t} = F(t)$, on peut remplacer $F(t)$ par son échantillonnée $F_{\text{ech}}(t) = T\sum_n F(nT)\,\delta(t-nT)$. On obtient alors, en intégrant entre les instants $nT-\epsilon$ et $nT+\epsilon$, la **discrétisation** $m\,\dfrac{v_n - v_{n-1}}{T} = F(nT)$ de l'équation de départ ($v_n = v(nT+\epsilon)$).

Les exemples ci-dessus permettent de comprendre comment résoudre les équations

$$\dot v + \frac{v}{\tau} = \delta(t)\quad(1) \qquad \text{et} \qquad \ddot x + 2\lambda\dot x + \omega_0^2 x = \delta(t)\quad(2) \qquad (\lambda < \omega_0)$$

avec conditions initiales nulles à $t = -\infty$ (particule en milieu visqueux et oscillateur amorti au repos soumis à une percussion à $t = 0$). Comme $\delta(t) = 0$ pour $t \neq 0$, on a $v(t)$ ou $x(t) = 0$ pour $t < 0$, et a priori $v(t) = Ae^{-\frac{t}{\tau}}$ ou $x(t) = A\,e^{-\lambda t}\cos(\omega t + \varphi)$ (avec $\omega = \sqrt{\omega_0^2 - \lambda^2}$) pour $t > 0$. L'effet de $\delta(t)$ étant simplement une variation brutale de vitesse de 1 à $t = 0$, on a $v(0^+) = \dot x(0^+) = 1$, d'où les solutions (**réponse impulsionnelle**) :

$$R(t) = H(t)\,e^{-\frac{t}{\tau}}\quad(1) \qquad ; \qquad R(t) = H(t)\,\frac{\sin\omega t}{\omega}\,e^{-\lambda t}\quad(2)\ .$$

REMARQUE : le saut de vitesse étant fini il ne peut y avoir de déplacement instantané à $t = 0$ induit par $\delta(t)$.

5.2.3 Analyse de Dirac ; réponse impulsionnelle ; convolution ; filtrage

La formule (conséquence immédiate des propriétés de δ)

$$f(x) = \int_{-\infty}^{\infty} f(a)\,\delta(x-a)\,\mathrm{d}a = [f * \delta](x)\ ,$$

montre, d'une part que $\delta(x)$ est l'**unité de la convolution**, et d'autre part que toute fonction $f(x)$ peut s'interpréter comme somme "continue" d'impulsions $f(a)\,\delta(x-a)$, contribution à $f(x)$ de chaque point a de son support. Cette décomposition de $f(x)$ sur la base des fonctions $\delta(x-a)$, limite continue de l'échantillonnage, constitue l'**analyse de Dirac** de $f(x)$.

▷ Réponse impulsionnelle ; convolution

Considérons une opération linéaire et stationnaire et soit $R(x)$ (réponse impulsionnelle) le résultat de cette opération sur $\delta(x)$. La stationnarité et la linéarité entraînent qu'à la fonction (de x) $f(a)\delta(x-a)$ correspond le résultat $f(a)R(x-a)$ et, en sommant sur a, que l'opération sur $f(x)$ s'écrit sous la forme d'une convolution :

$$f(x) \to \int_{-\infty}^{\infty} f(a)\,R(x-a)\,\mathrm{d}a = [R * f](x)\ .$$

EXEMPLE 1. **Circuit RC**. Soit $e(t)$ ("entrée") la tension aux bornes du générateur et $s(t)$ ("sortie") celle aux bornes de C. On déduit des formules (1) ci-dessus

$$\tau\dot{s} + s = e \Longrightarrow s(t) = \int_{-\infty}^{t} \frac{1}{\tau}\, e^{-\frac{t-t'}{\tau}}\, e(t')\,\mathrm{d}t'\ ,$$

où $\tau = RC$ et $R(t) = H(t)\,\dfrac{1}{\tau}\,e^{-\frac{t}{\tau}}$. Si τ est grand devant le temps caractéristique de variation de $e(t)$, alors $s(t) \simeq \displaystyle\int_{-\infty}^{t} \frac{1}{\tau}\,e(t')\,\mathrm{d}t'$ (circuit **intégrateur**). Si τ est petit l'exponentielle est concentrée autour de $t' = t$ et $s(t) \simeq e(t)$; en prenant alors comme sortie la tension s_R aux bornes de la résistance, on obtient $s_R(t) = \tau\dot{s}(t) \simeq \tau\dot{e}(t)$ (circuit **dérivateur**).

EXEMPLE 2. **Principe d'Huygens Fresnel**. Si $f(x,y,0)$ est l'amplitude complexe d'une onde en tout point m du plan $z=0$ (à la sortie d'un écran), la formule d'Huygens-Fresnel (*cf.* section 2.5.2) dit qu'en tout M du plan z (près de l'axe) :

$$f(X,Y,z) = \frac{1}{i\lambda z}\, e^{i2\pi\frac{z}{\lambda}} \iint f(x,y,0)\, e^{i\pi\frac{(X-x)^2+(Y-y)^2}{z\lambda}}\,\mathrm{d}x\,\mathrm{d}y\ .$$

La propagation de $z=0$ à $z>0$ se traduit donc par une opération de convolution (à deux dimensions) dont la réponse impulsionnelle est l'onde sphérique

$$R_z(x,y) = \frac{1}{i\lambda z}\, e^{i\frac{2\pi}{\lambda}\left(z+\frac{x^2+y^2}{2z}\right)} \simeq \frac{1}{i\lambda r}\, e^{i\frac{2\pi}{\lambda}\,r}$$

émise par une source ponctuelle en O ($r = \sqrt{x^2+y^2+z^2}$).

EXEMPLE 3. **Electrostatique**. La relation entre une distribution volumique de charges $\rho(\vec{r})$ et le potentiel qu'elle crée $V(\vec{r})$ est elle aussi donnée par un produit de convolution (à trois dimensions)

$$V(\vec{r}) = \tfrac{1}{4\pi\epsilon_0} \iiint \tfrac{\rho(\vec{r}_0)}{|\vec{r}-\vec{r}_0|}\, \mathrm{d}^3 r_0$$

dont la réponse impulsionnelle $\frac{1}{4\pi\epsilon_0 r}$ est le potentiel créé par une charge ponctuelle unité en O.

EXEMPLE 4. **Optique incohérente**. La répartition d'intensité, dans le plan (x', y'), de l'image d'un objet incohérent dans le plan (x, y) s'écrit $I'(x', y') = \iint R(x' - \gamma x, y' - \gamma y)\, I(x, y)\, \mathrm{d}x\, \mathrm{d}y$, où $R(x', y')$ est l'intensité de la tache de diffraction associée à une source ponctuelle $\delta(x)\delta(y)$ sur l'axe optique ; γ est le grandissement.

EXEMPLE 5. **Signaux numériques** (*cf.* section 2.4.2). Le signal $\{R_n\}$ joue le rôle de la réponse impulsionnelle.

▷ Filtrage

La convolution appliquée à une exponentielle imaginaire donne (calcul déjà vu à la section 2.4.2) :

$$R * Ae^{i2\pi\nu t} = \int_{-\infty}^{\infty} R(t-t')\, Ae^{i2\pi\nu t'}\, \mathrm{d}t' = \hat{R}(\nu)\, Ae^{i2\pi\nu t} \qquad (\hat{R}(\nu) = \int_{-\infty}^{\infty} e^{-i2\pi\nu t}\, R(t)\, \mathrm{d}t)\ .$$

Elle a donc pour effet de modifier le module et la phase de l'amplitude complexe du signal de fréquence ν. Ce résultat est très important quand on sait que tout signal peut s'écrire comme une somme d'exponentielles complexes (analyse de Fourier ; *cf.* section 5.3.1). La linéarité entraîne :

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\nu)\, e^{i2\pi\nu t}\, \mathrm{d}\nu \quad \longrightarrow \quad [R * f](t) = \int_{-\infty}^{\infty} \hat{f}(\nu)\, \hat{R}(\nu)\, e^{i2\pi\nu t}\, \mathrm{d}\nu\ .$$

L'opération de convolution revient à multiplier chaque composante de Fourier $\hat{f}(\nu)$ du signal par $\hat{R}(\nu)$, **fonction de transfert** de l'opération de filtrage. Suivant que $|\hat{R}(\nu)|$ est grand surtout pour ν petit, pour ν grand, ou pour ν dans un intervalle, on a affaire à un filtre passe bas, passe haut, ou passe bande. Si $\hat{R}(\nu) = e^{i\varphi(\nu)}$ on parle de filtre de phase ; $\hat{R}(\nu) = e^{-i2\pi\nu\tau}$ pour une ligne à retard $f(t) \to f(t-\tau)$ car : $e^{i2\pi\nu t} \to e^{i2\pi\nu(t-\tau)}$.

REMARQUE. Il est important de noter qu'en physique on a souvent un accès plus direct à la fonction de transfert $\hat{R}(\nu)$ qu'à la réponse impulsionnelle $R(t)$. Par exemple pour un **circuit RC**, $\hat{R}(\nu) = \displaystyle\int_{-\infty}^{\infty} e^{-i2\pi\nu t}\, H(t)\, \frac{1}{\tau}\, e^{-\frac{t}{\tau}}\, \mathrm{d}t = (1 + i2\pi\nu\tau)^{-1}$ n'est rien d'autre que le coefficient de $e^{i2\pi\nu t}$ dans la solution stationnaire de $\tau\, \dot{s} + s = e^{i2\pi\nu t}$ (*cf.* section 2.4.2). $\hat{R}(\nu)$ correspond à un filtre passe bas du premier ordre, intégrateur si $2\pi\nu\tau = \omega\tau \gg 1$.

5.3 TRANSFORMATION DE FOURIER ; ANALYSE DE FOURIER

L'analyse de Fourier s'identifie en physique à la "volonté" de décomposer les fonctions sur la base, discrète ou continue, constituée par les exponentielles imaginaires, fonctions caractérisées par une fréquence "pure". Comme on l'a vu ci-dessus elle est bien adaptée à la physique linéaire et stationnaire et à la notion de filtrage. Historiquement introduite pour les fonctions périodiques à propos de l'étude de l'équation de la chaleur (Fourier 1822), elle s'applique à une classe très large de fonctions, aux distributions (donc à δ), et sert aussi à distinguer les signaux stationnaires chaotiques des signaux quasi-périodiques (*cf.* section 5.3.3). Les formules seront écrites pour des signaux $f(t)$ de **transformée de Fourier (T.F.)** $\hat{f}(\nu)$.

5.3.1 Décomposition de Fourier ; spectre d'un signal ; formule de Poisson

▷ Formules de base et justifications

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(\nu)\, e^{i2\pi\nu t}\, \mathrm{d}\nu \quad \text{avec} \quad \hat{f}(\nu) = \int_{-\infty}^{\infty} f(t)\, e^{-i2\pi\nu t}\, \mathrm{d}t \ .$$

$$f_{\mathrm{per}}(t) = \sum_{n=-\infty}^{\infty} f_n\, e^{i2\pi\frac{n}{T}t} \ \text{avec} \ f_n = \frac{1}{T}\int_{t_0-\frac{T}{2}}^{t_0+\frac{T}{2}} f_{\mathrm{per}}(t)\, e^{-i2\pi\frac{n}{T}t}\, \mathrm{d}t \overset{\text{déf}}{=} < f_{\mathrm{per}}(t)\, e^{-i2\pi\frac{n}{T}t} > .$$

La première formule est une conséquence directe de la **décomposition de Fourier de la fonction de Dirac** :

$$\delta(t) = \int_{-\infty}^{\infty} e^{i2\pi\nu t}\, \mathrm{d}\nu \ .$$

JUSTIFICATION. La dernière égalité s'obtient en considérant la limite $\epsilon \to 0$ d'expressions telles que :

$$\int_{-1/2\epsilon}^{1/2\epsilon} e^{i2\pi\nu t}\, \mathrm{d}\nu = \frac{1}{\epsilon}\frac{\sin \pi t/\epsilon}{\pi t/\epsilon} \ \text{ ou } \ \int_{-\infty}^{\infty} e^{-2\pi\epsilon|\nu|}\, e^{i2\pi\nu t}\, \mathrm{d}\nu = \frac{\epsilon}{\pi(t^2+\epsilon^2)}$$

(*cf.*définition de $\delta(t)$ à la section 5.2.2). Comme $f = f * \delta$, on en déduit :

$$f(t) = \int f(t') \left[\int e^{i2\pi\nu(t-t')}\, \mathrm{d}\nu\right] \mathrm{d}t' = \int \hat{f}(\nu)\, e^{i2\pi\nu t}\, \mathrm{d}\nu \ .$$

REMARQUES : 1) La correspondance $f(t) \Longleftrightarrow \hat{f}(\nu)$ est biunivoque ; le passage de $\hat{f}(\nu)$ à $f(t)$ s'appelle la T.F. inverse. 2) La dimension de $\hat{f}$ est $[\hat{f}] = [f][t]$.

La seconde formule concerne seulement les fonctions périodiques. L'expression donnant les coefficients de la série de Fourier f_n (moyenne temporelle sur une période T), découle des relations $< e^{i2\pi\frac{p}{T}t}\, e^{-i2\pi\frac{q}{T}t} >= 0$ si $p \neq q$ et 1 si $p = q$ (en particulier f_0 est la valeur moyenne $< f(t) >$). Une formule équivalente, dite **formule de Poisson**, est :

$$f_{\mathrm{per}}(t) = \sum_{n=-\infty}^{\infty} f(t - nT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \hat{f}\left(\frac{n}{T}\right)\, e^{i2\pi\frac{n}{T}t} \ .$$

JUSTIFICATION. Cette formule est une conséquence directe de la relation facile à vérifier

$$f_{\mathrm{per}}(t) = f * \textstyle\sum_{n=-\infty}^{\infty} \delta(t - nT) \ ,$$

et de la **décomposition de Fourier du peigne de Dirac** :

$$\textstyle\sum_{n=-\infty}^{\infty} \delta(t - nT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} e^{i2\pi\frac{n}{T}t} \ .$$

Cette dernière s'obtient en considérant la limite $\epsilon \to 0$ de $\sum_{n=-\infty}^{\infty} e^{-\epsilon|n|}\, e^{in\varphi}$, ou la limite $N \to \infty$ de $\sum_{n=-N}^{N} e^{in\varphi}$ (*cf.* figures 27 et 28 du chapitre 2).

REMARQUE. Dans la limite $T \to \infty$, la formule de Poisson redonne $f(t) = \int_{-\infty}^{\infty} \hat{f}(\nu)\, e^{i2\pi\nu t}\, \mathrm{d}\nu$: le membre de gauche devient $f(t)$ (les $f(t - nT)$, $n \neq 0$, "partant à l'infini"), et la somme de Riemann du second membre se transforme en intégrale.

▷ Spectre d'un signal

C'est l'ensemble des fréquences ν impliquées dans la décomposition de Fourier du signal. Le spectre peut être continu (*cf.* exemples ci-dessous) ou discret (par exemple le signal $f(t) = e^{i2\pi\nu_0 t}$ obtenu pour $\hat{f}(\nu) = \delta(\nu - \nu_0)$). Dans le premier cas il existe en général un lien entre les propriétés de régularité de $\hat{f}(\nu)$ et la décroissance plus ou moins rapide de $f(t)$ lorsque $t \to \pm\infty$ (et réciproquement ; *cf.* section 5.3.2. "dérivation").
Les fonctions périodiques ont un spectre discret dont les fréquences $\dfrac{n}{T}$ (**harmoniques** d'ordre n) sont multiples de la fréquence fondamentale $\dfrac{1}{T}$. Cette fréquence peut ne pas apparaitre, par exemple $f_{\rm per}(t) = \cos 3t + \cos 5t$ de période 2π ne contient que les harmoniques ± 3 et ± 5). Remarque : traditionnellement on utilise plutôt la décomposition qui fait intervenir les coefficients de Fourier que la formule de Poisson. Cependant cette dernière est non seulement très utile, puisqu'on obtient en une seule fois tous les f_n à partir de $\hat{f}(\nu)$, mais aussi plus riche de sens physique : elle montre que périodiser une fonction f à la période T équivaut à échantillonner sa T.F. à la période $\dfrac{1}{T}$ c.a.d. consiste à ne retenir dans son spectre que les fréquences $\nu_n = \dfrac{n}{T}$.

▷ Fonctions réelles

Si $f(t)$ est réelle $\hat{f}(-\nu) = \overline{\hat{f}}(\nu)$ ou $f_{-n} = \overline{f_n}$: toute l'information concernant le signal est contenue dans la partie $\nu \geq 0$ du spectre. On préfère alors souvent écrire les développements à l'aide de fonctions réelles. Par exemple

$$f_{\rm per}(t) = a_0 + \sum_{n\geq 1} \left(a_n \cos 2\pi \frac{n}{T} t + b_n \sin 2\pi \frac{n}{T} t \right) ,$$

avec $a_0 =< f_{\rm per}(t) >= f_0$, $a_n = 2 < f_{\rm per}(t) \cos 2\pi \frac{n}{T} t >$ et $b_n = 2 < f_{\rm per}(t) \sin 2\pi \frac{n}{T} t >$, donc $2f_n = a_n - ib_n$ pour $n \geq 1$; $b_n = 0$ si $f_{\rm per}$ est paire et $a_n = 0$ si $f_{\rm per}$ est impaire.

▷ Exemples de T.F.

Parmi les exemples courants facilement calculables, citons la fonction **porte** de largeur τ centrée $\Pi\left(\dfrac{t}{\tau}\right) \overset{\rm TF}{\Longrightarrow} \tau \,{\rm sinc}\pi\tau\nu$ et $\nu_0 \,{\rm sinc}\pi\nu_0 t \overset{\rm TF^{-1}}{\Longleftarrow} \Pi\left(\dfrac{\nu}{\nu_0}\right)$ (fonction de transfert d'un **filtre passe bas idéal**) ; $\delta(t) \overset{\rm TF}{\Longrightarrow} 1$ (donc dans la fonction δ toutes les fréquences apparaissent avec le même poids), et $1 \overset{\rm TF}{\Longrightarrow} \delta(\nu)$, qui sont des limites des cas précédents ; $\exp\left(-\dfrac{|t|}{\tau}\right) \overset{\rm TF}{\Longrightarrow} \dfrac{2\tau}{1+(2\pi\tau\nu)^2}$; $H(t)\, e^{-\frac{t}{\tau}} \overset{\rm TF}{\Longrightarrow} \dfrac{\tau}{1+i2\pi\tau\nu}$ (filtre passe bas du premier ordre). Un autre exemple important dans la pratique est celui de la fonction **gaussienne** $\exp\left(-\pi t^2\right) \overset{\rm TF}{\Longrightarrow} \exp\left(-\pi\nu^2\right)$, et aussi en optique $\dfrac{1}{\sqrt{i}} \exp(i\pi x^2) \overset{\rm TF}{\Longrightarrow} \exp(-i\pi\sigma^2)$.

▷ Exemple de fonction périodique (figure 18)

Les coefficients f_n de la fonction créneaux s'obtiennent en prenant la moyenne sur $\left[-\dfrac{T}{2}, \dfrac{T}{2}\right]$: $f_n = \dfrac{1}{T} \displaystyle\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} A\, e^{-i2\pi\frac{n}{T}t}\, {\rm d}t$. On en déduit le développement :

$$A \sum_{n=-\infty}^{\infty} \Pi\left(\frac{t-nT}{\tau}\right) = A\,\frac{\tau}{T} \sum_{n=-\infty}^{\infty} e^{i2\pi\frac{n}{T}t} \,{\rm sinc}\pi\tau\frac{n}{T} \ .$$

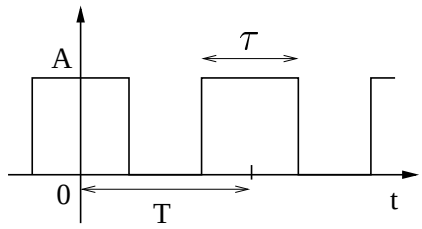


FIGURE 18

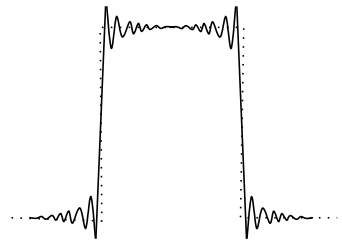


FIGURE 19

On vérifie sur cet exemple que $f_n = \dfrac{1}{T}\,\hat{f}\left(\dfrac{n}{T}\right)$ où $f(t) = A\,\Pi\left(\dfrac{t}{\tau}\right)$ (formule de Poisson). Dans la limite $\tau \to 0$, $A \to \infty$ avec $A\tau = 1$, on retrouve la relation $\displaystyle\sum_{n=-\infty}^{\infty} \delta(t - nT) = \frac{1}{T}\sum_{n=-\infty}^{\infty} e^{i2\pi\frac{n}{T}t}$. Pour $\tau = \dfrac{T}{2}$, on remarque que les fréquences $\nu_{2p} = \dfrac{2p}{T}$ $(p \neq 0)$ disparaissent et que pour $t = \dfrac{T}{4}$ la série vaut $\dfrac{A}{2}$, c.a.d. prend la valeur moyenne des deux valeurs situées de part et d'autre de la discontinuité. Ce dernier résultat est général dans les développements de fonctions discontinues ; il s'accompagne au niveau numérique, quand on tronque la série, de grandes oscillations près de la discontinuité (**phénomène de Gibbs**) ; la figure 19 correspond aux 15 premiers termes pour la fonction créneaux.

5.3.2 Propriétés : dualité temps fréquence

En choisissant bien une des expressions donnant $f(t)$ ou $\hat{f}(\nu)$, on déduit en général sans calcul les propriétés suivantes.

- **Linéarité** : $\lambda f + \mu g \overset{\text{TF}}{\Longrightarrow} \lambda\hat{f} + \mu\hat{g}$.

- **Aires** : $\hat{f}(0) = \displaystyle\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t \quad ; \quad f(0) = \int_{-\infty}^{\infty} \hat{f}(\nu)\,\mathrm{d}\nu$.

- **Réalité** : $f(t)$ réelle $\Longleftrightarrow \hat{f}(-\nu) = \overline{\hat{f}(\nu)}$.

- **Dilatation** : $f\left(\dfrac{t}{k}\right) \overset{\text{TF}}{\Longrightarrow} k\,\hat{f}(k\nu)$.

Donc plus une fonction varie vite (resp. lentement), plus les hautes (resp. basses) fréquences doivent être présentes dans son spectre. Pour $k \to 0$ (resp. $k \to \infty$) on retrouve $\delta(t) \overset{\text{TF}}{\Longrightarrow} 1$ (resp. $1 \overset{\text{TF}}{\Longrightarrow} \delta(\nu)$). Cette dualité entre les échelles de temps et de fréquence peut se traduire qualitativement par $\Delta t\,\Delta\omega \simeq 1$, formule analogue aux relations de Heisenberg $\Delta t\,\Delta E \simeq \hbar$ ou $\Delta x\,\Delta p_x \simeq \hbar$ en mécanique quantique.

- **Parité** : $f(-t) \overset{\text{TF}}{\Longrightarrow} \hat{f}(-\nu)$.

Donc si f réelle est paire (resp. impaire), $\hat{f}$ est réelle paire (resp. imaginaire pure et impaire).

- **Translation et multiplication par une exponentielle** :

$$f(t-\tau) \overset{\text{TF}}{\Longrightarrow} e^{-2i\pi\tau\nu}\,\hat{f}(\nu) \quad \text{et} \quad e^{2i\pi\nu_0 t}\,f(t) \overset{\text{TF}}{\Longrightarrow} \hat{f}(\nu-\nu_0)\ .$$

Si l'instant d'émission d'un signal est décalé dans le temps, ceci ne se traduit donc que par un facteur de phase au niveau de sa T.F. La **modulation d'amplitude**

$$x(t) \to f(t) = A\left(1+mx(t)\right)\,\cos 2\pi\nu_0 t$$

a pour but de transférer le spectre basse fréquence de $x(t)$ ($|\nu| \in [0,B]$) dans celui de haute fréquence de $f(t)$ ($|\nu| \in [\nu_0-B, \nu_0+B]$) qui permet les télécommunications. Après transmission on peut démoduler par une opération de **détection synchrone** $f(t) \to f(t)\,\cos 2\pi\nu_0 t$ (spectre $|\nu| \in [0,B]$ et $|\nu| \in [2\nu_0-B, 2\nu_0+B]$ car $\cos^2 2\pi\nu_0 t = \dfrac{1}{2}\,(1+\cos 4\pi\nu_0 t)$), suivie d'un filtrage passe-bas qui redonne $x(t)$.

- **Dérivation et multiplication par une puissance** :

$$\frac{\mathrm{d}^n f(t)}{\mathrm{d}t^n} \overset{\text{TF}}{\Longrightarrow} (i2\pi\nu)^n\,\hat{f}(\nu) \quad ; \quad (-2i\pi t)^n\,f(t) \overset{\text{TF}}{\Longrightarrow} \frac{\mathrm{d}^n \hat{f}(\nu)}{\mathrm{d}\nu^n}\ .$$

Si un signal a une **discontinuité** finie en t_0, sa dérivée contient un $\delta(t-t_0)$ et $\hat{f}(\nu)$ se comporte comme $|\nu|^{-1}$ à l'infini (*cf.* fonction porte) ; un comportement en $|\nu|^{-2}$ traduit une discontinuité de f' (*cf.* $\exp\left(-\dfrac{|t|}{\tau}\right)$)... : plus f est régulière plus $\hat{f}$ décroît vite à l'infini (et réciproquement).

- **Convolution et produit** :

$$[f*g]\,(t) \overset{\text{TF}}{\Longrightarrow} \hat{f}(\nu)\,\hat{g}(\nu) \quad ; \quad f(t)\,g(t) \overset{\text{TF}}{\Longrightarrow} \left[\hat{f}*\hat{g}\right](\nu)\ .$$

La première relation se démontre en écrivant :

$$[f*g]\,(t) = \int_{-\infty}^{\infty} f(t')\left[\int_{-\infty}^{\infty} \hat{g}(\nu)\ e^{i2\pi\nu(t-t')}\,\mathrm{d}\nu\right]\mathrm{d}t' = \int_{-\infty}^{\infty} \hat{f}(\nu)\,\hat{g}(\nu)\,e^{i2\pi\nu t}\,\mathrm{d}\nu\ .$$

Sa traduction en termes de filtrage a déjà été vue à la section 5.2.3. Pour la seconde on vérifie $\left[\hat{f}*\hat{g}\right](\nu) = \int_{-\infty}^{\infty} \hat{f}(\nu')\left[\int_{-\infty}^{\infty} g(t)\ e^{-i2\pi(\nu-\nu')t}\,\mathrm{d}t\right]\mathrm{d}\nu' = \int_{-\infty}^{\infty} f(t)\,g(t)$ $e^{-i2\pi\nu t}\,\mathrm{d}t$. Si $g(t) = \Pi\left(\dfrac{t}{T}\right)$ (troncature de $f(t)$), la convolution de $\hat{f}(\nu)$ par $\hat{g}(\nu) = T\,\mathrm{sinc}\pi\nu T$ conduit à un élargissement du spectre de f.

- **Causalité** : l'idée de causalité ("l'effet ne précède pas la cause") se traduit sur la réponse impulsionnelle par $R(t) = 0$ pour $t < 0$, soit

$$R(t) = \Theta(t)\,R(t) \ \text{ avec } \ \Theta(t) = 1 \text{ pour } t > 0 \text{ et } -1 \text{ pour } t < 0, \ \text{ ou } \ \widehat{R}(\nu) = [\widehat{\Theta} * \widehat{R}](\nu)\ .$$

Comme $\Theta(t)$ est réelle impaire, $\hat{\Theta}(\nu)$ est imaginaire pure ; ceci entraîne l'existence de relations (**Kramers-Kronig**) entre les parties réelle et imaginaire de la fonction de transfert $\widehat{R}(\nu)$. Un exemple est donné par la réponse $\vec{P}(t)$, polarisation diélectrique d'un milieu non parfait, à un champ électrique $\vec{E}(t)$ et la fonction de transfert associée $\epsilon(\omega) - \epsilon_0$; la **permittivité complexe** $\epsilon(\omega)$ étant reliée à l'indice complexe par $\epsilon(\omega) = n^2(\omega)\,\epsilon_0$, tout milieu dispersif ($\Re e\,\epsilon(\omega)$ fonction dépendant de ω) est en général aussi dissipatif ($\Im m\,\epsilon(\omega) \neq 0$), c.a.d. le siège de pertes d'énergie.

REMARQUES. 1) Un **filtre passe bas idéal** ($R(t) = \nu_0 \operatorname{sinc}\pi\nu_0 t$) n'est pas causal. 2) Techniquement $\widehat{\Theta}(\nu)$ est la limite pour $\epsilon \to 0$ de $\int_{-\infty}^{\infty} \Theta(t)\, e^{-\epsilon|t|}\, e^{-i2\pi\nu t}\, \mathrm{d}t = \frac{4\pi\nu}{i(\epsilon^2+4\pi^2\nu^2)}$ (régularisation de $\frac{1}{i\pi\nu}$ près de $\nu = 0$).

- **Autocorrélation, densité spectrale d'énergie**. Un calcul semblable à celui fait pour la convolution conduit à la décomposition spectrale d'une fonction de corrélation

$$\Gamma_f(\tau) = \int_{-\infty}^{\infty} \overline{f}(t)\, f(t+\tau)\, \mathrm{d}t = \int_{-\infty}^{\infty} |\hat{f}(\nu)|^2\, e^{i2\pi\nu\tau} \mathrm{d}\nu \ ,$$

qui pour $\tau = 0$ donne la **formule de Parseval** :

$$\int_{-\infty}^{\infty} |f(t)|^2\, \mathrm{d}t = \int_{-\infty}^{\infty} |\hat{f}(\nu)|^2\, \mathrm{d}\nu \quad (= 2\int_0^{\infty} |\hat{f}(\nu)|^2\, \mathrm{d}\nu \text{ pour } f \text{ réelle}) \ .$$

Par exemple l'énergie $E = \int_{-\infty}^{\infty} |f(t)|^2\, \mathrm{d}t$ contenue dans un signal lumineux est la somme des contributions de chaque fréquence : composantes violette, bleue...rouge de son spectre obtenu avec un prisme. La densité spectrale d'énergie d'un signal réel $E(\nu) = 2|\hat{f}(\nu)|^2$ est donc reliée a la T.F. de Γ_f. Ce résultat est l'équivalent pour les signaux d'énergie finie du **théorème de Wiener-Khintchine** pour les signaux stationnaires (*cf.* sections 5.3.3 et 10.3.3).

- **Périodisation et échantillonnage**. On a déjà vu (formule de Poisson) que périodiser $f(t)$ conduit à échantillonner $\hat{f}(\nu)$; inversement échantillonner $f(t)$ conduit à périodiser $\hat{f}(\nu)$:

$$f_{\mathrm{ech}}(t) = T \sum_{n=-\infty}^{\infty} f(nT)\, \delta(t-nT) \ \overset{\mathrm{TF}}{\Longrightarrow}\ \widehat{f_{\mathrm{ech}}}(\nu) = \sum_{n=-\infty}^{\infty} \hat{f}\left(\nu - \frac{n}{T}\right) \ .$$

Cette relation se démontre en remarquant que $f_{\mathrm{ech}}(t)$ s'écrit aussi $f_{\mathrm{ech}}(t) = T \sum_n f(t)$ $\delta(t-nT) = f(t) \sum_n e^{i2\pi\frac{n}{T}t}$. Application : si $f(t)$ a son spectre contenu dans $[-B, B]$, et si l'échantillonnage est assez serré ($T < \dfrac{1}{2B}$ **condition de Shannon**), les graphes des fonctions $\hat{f}\left(\nu - \dfrac{n}{T}\right)$ ne se recouvrent pas (figure 20). On peut alors reconstruire $f(t)$ à partir de $f_{\mathrm{ech}}(t)$ en effectuant une opération de filtrage dont la fonction de transfert $\hat{R}(\nu)$ vaut 1 sur le support de $\hat{f}(\nu)$ et 0 sur ceux des fonctions $\hat{f}(\left(\nu - \dfrac{n}{T}\right)$ $(n \neq 0)$. On en déduit :

$$f(t) = T \sum_{n=-\infty}^{\infty} R(t-nT)\, f(nT) \ ;$$

par exemple pour $\hat{R}(\nu) = \Pi\left(\dfrac{\nu}{B}\right)$ on a $R(t) = 2B \operatorname{sinc} 2\pi B t$, mais en resserrant l'échantillonnage et en prenant $\widehat{R}(\nu)$ plus régulière, on peut assurer une convergence plus rapide de la somme sur n.

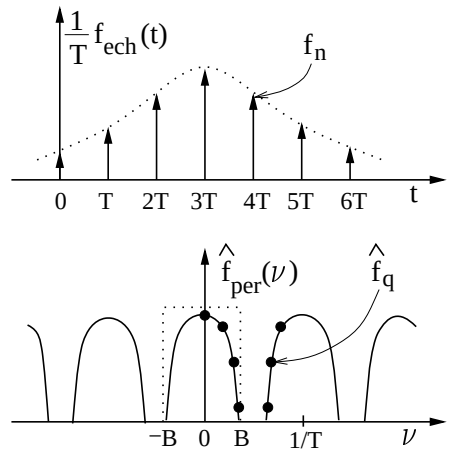


FIGURE 20

REMARQUE. Le facteur T dans l'expression de $f_{\rm ech}(t)$ est nécessaire pour que $f_{\rm ech}$ soit dimensionnellement homogène à f et pour que $f_{\rm ech}(t)$ et $\widehat{f_{\rm ech}}(\nu)$ redonnent $f(t)$ et $\hat{f}(\nu)$ dans la limite "continue" $T \to 0$.

- **Transformée de Fourier discrète (T.F.D)**. Le traitement des signaux par les calculateurs nécessite le passage de l'analogique (signaux continus) au numérique (signaux discrétisés). La T.F.D. (et son inverse) relie N échantillons $\hat{f}_q$ représentatifs de $\hat{f}(\nu)$ à N échantillons $f_n = f(nT)$ représentatifs de $f(t)$ par :

$$\hat{f}_q = \tfrac{1}{N} \sum_{n=0}^{N-1} f_n\, e^{-i2\pi \frac{n}{N} q} \quad ; \quad f_n = \sum_{q=0}^{N-1} \hat{f}_q\, e^{i2\pi \frac{q}{N} n} \ .$$

Pour vérifier que les $\hat{f}_q$ ainsi définis conviennent, on écrit $\widehat{f_{\rm ech}}(\nu) = T \sum_{n=-\infty}^{\infty} f(nT)\, e^{-i2\pi\nu nT}$ (calcul direct de la T.F. de $f_{\rm ech}(t)$). Si $f(t)$ est convenablement centrée et négligeable en dehors de $[0, NT[$, la fonction $\widehat{f_{\rm N-ech}}(\nu) = T \sum_{n=0}^{N-1} f_n\, e^{-i2\pi\nu nT}$ est proche de $\widehat{f_{\rm ech}}(\nu)$. Si de plus l'échantillonnage de f est assez serré, les valeurs de $\widehat{f_{\rm ech}}(\nu)$ pour $\nu \in \left[0, \frac{1}{T}\right]$ sont égales (modulo un réarrangement) à celles de $\hat{f}(\nu)$ pour $\nu \in \left[-\frac{1}{2T}, \frac{1}{2T}\right]$ (figure 20). L'ensemble des valeurs $\widehat{f_{\rm N-ech}}\left(\frac{q}{NT}\right) = T \sum_{n=0}^{N-1} f_n\, e^{-i2\pi \frac{q}{N} n} = NT \hat{f}_q$ $(q = 0, 1, \cdots, N-1)$ est donc bien représentatif de $\hat{f}(\nu)$. L'intérêt numérique de la T.F.D. est que pour $N = 2^p$ il existe un algorithme (**T.F. rapide ou F.F.T.** en anglais) qui ramène les N^2 (en gros) opérations nécessaires au calcul des $\hat{f}_q$ à seulement $\frac{N}{2}(p-1)$. Un autre intérêt est l'existence, pour chaque propriété de dualité écrite ci-dessus pour $f(t)$ et $\hat{f}(\nu)$, d'une propriété analogue pour les suites de nombres f_n et $\hat{f}_q$ (*cf.* ouvrages spécialisés).

5.3.3 Transformée de Laplace

Elle est donnée, pour les signaux définis sur $[0, \infty]$, par

$$f(t) \to F(p) = \int_0^\infty e^{-pt}\, f(t)\, {\rm d}t \ ,$$

et vérifie les propriétés (faciles à établir) :

$$\delta(t) \to 1,\ e^{-at} \to (p+a)^{-1},\ \cos\omega t \to p\,(p^2+\omega^2)^{-1},\ \sin\omega t \to \omega\,(p^2+\omega^2)^{-1} \ ,$$

$$\frac{\rm d}{{\rm d}t} f(t) \to p\,F(p) - f(0),\ \frac{{\rm d}^2}{{\rm d}t^2}\, f(t) \to p^2\,F(p) - \frac{{\rm d}f}{{\rm d}t}(0) - p\,f(0) \ ,$$

$$\big[f * g\big](t) = \int_0^t f(s)\, g(t-s)\, \mathrm{d}s \to F(p)\, G(p),\ e^{-\lambda t} f(t) \to F(p+\lambda) \quad (\Re e\lambda \geq 0)\ .$$

Elle a l'intérêt de prendre en compte les C.I. pour la résolution des E.D. ; par exemple de $\ddot{x} + \omega^2 x = f(t) \Longrightarrow X(p) = (p^2+\omega^2)^{-1}\big(F(p) + \dot{x}(0) + px(0)\big)$, on déduit $x(t) = x(0)\,\cos\omega t + \dot{x}(0)\,\dfrac{\sin\omega t}{\omega} + \displaystyle\int_0^t \frac{\sin\omega(t-s)}{\omega}\, f(s)\,\mathrm{d}s$ (*cf.* aussi section 6.2.1).

5.3.4 Signaux stationnaires ; signaux chaotiques ; langage probabiliste

▷ Signaux stationnaires ; détection synchrone

Ce sont, idéalement, des signaux pour lesquels toutes les **moyennes temporelles**

$$< \cdots >= \lim_{T\to\infty}\ \frac{1}{T}\ \int_{t_0-\frac{T}{2}}^{t_0+\frac{T}{2}} (\cdots)\,\mathrm{d}t$$

ont un sens (indépendant de t_0), par exemple pour un signal réel les moments $< (f(t))^n >$ ou la fonction de corrélation $\Gamma(\tau) =< f(t)\, f(t+\tau) >$. De tels signaux ont une énergie E infinie mais une "puissance moyenne" $\Gamma(0)$ finie. Parmi eux il y a les signaux périodiques mais aussi plus généralement les **signaux quasi-périodiques** qui, par définition, ont un **spectre discret** sur lequel est concentrée leur puissance :

$$f_{\mathrm{q-per}}(t) = A_0 + \sum_{\nu_n>0} A_n\, \cos(2\pi\nu_n t + \varphi_n) \quad ; \quad \Gamma(\tau) = A_0^2 + \sum_n\ \frac{A_n^2}{2}\, \cos 2\pi\nu_n\tau\ .$$

Le signal est périodique si les fréquences sont commensurables, par exemple $\cos 2t + \cos(3,14\,t)$ mais pas $\cos 2t + \cos\pi t$ qui est quasi-périodique ; cet exemple montre qu'il n'est pas toujours facile de distinguer pratiquement ces deux types de signaux.

Pour analyser un signal stationnaire $f(t)$, on peut le multiplier par une fonction sinusoïdale et effectuer la moyenne temporelle du produit ; on détecte ainsi la présence de la fréquence correspondante dans le spectre de $f(t)$ (méthode de **détection synchrone**) : si $f_{\nu_0} =< f(t)\, e^{-i2\pi\nu_0 t} >\neq 0$ on a $f(t) = f_{\nu_0}\, e^{i2\pi\nu_0 t} + g(t)$ avec $< g(t)\, e^{-i2\pi\nu_0 t} >= 0$. On peut (en principe) en explorant ainsi toutes les fréquences extraire la partie quasi-périodique d'un signal.

▷ Signaux chaotiques réels

Ces signaux apparaissent dans tous les domaines de la physique (fluctuations microscopiques, émission de la lumière par les atomes, télécommunications...). Tantôt ils jouent un rôle fondamental dans la connaissance des phénomènes (mouvement brownien, largeur de raie...), tantôt ils interviennent comme un "bruit" et on cherche à les éliminer (télécommunications...). Du point de vue de l'analyse de Fourier ce sont les signaux tels que $< f(t)\, e^{-i2\pi\nu t} >= 0$ quel que soit ν : ils sont orthogonaux aux $e^{i2\pi\nu t}$ et se caractérisent par un **spectre continu** (puissance concentrée sur aucune fréquence particulière). Leur fonction de corrélation s'écrit (**Théorème de Wiener-Khintchine**) :

$$\Gamma(\tau) = \int_0^\infty I(\nu)\, \cos 2\pi\nu\tau\, \mathrm{d}\nu \quad \text{avec} \quad I(\nu) = \lim_{T_0\to\infty}\ \frac{2}{T_0}\ \Big|\widehat{f_{T_0}}(\nu)\Big|^2\ .$$

Elle est une version "continue en fréquences" de celle des signaux quasi-périodiques.

DÉMONSTRATION : soit $f_{T_0}(t)$ le signal tronqué sur un intervalle de largeur T_0 et $\widehat{f_{T_0}}(\nu)$ sa T.F., on a $\Gamma(\tau) = \lim_{T_0\to\infty} \frac{1}{T_0}\int_{-\infty}^{\infty} f_{T_0}(t) f_{T_0}(t+\tau)\,\mathrm{d}t = \lim_{T_0\to\infty} \frac{1}{T_0}\int_{-\infty}^{\infty} \left|\widehat{f_{T_0}}(\nu)\right|^2 e^{i2\pi\nu\tau}\,\mathrm{d}\nu$ d'où l'expression de $I(\nu)$ pour un signal réel ($\hat{f}(\nu) = \overline{\hat{f}}(-\nu)$).

$\Gamma(\tau)$ a la propriété importante de tendre vers zéro lorsque $\tau \to \infty$ (ce qui n'est pas le cas pour un spectre discret, par exemple $\Gamma(\tau) = \cos 2\pi\nu_0\tau$ pour $I(\nu) = \delta(\nu - \nu_0)$). On appelle **temps de corrélation** τ_c une valeur typique de τ telle que $\Gamma(\tau) \simeq 0$ si $\tau > \tau_c$ (en général $\tau_c \simeq \frac{1}{\Delta\nu}$ où $\Delta\nu$ est la dispersion du spectre du signal). En optique τ_c est un temps de cohérence. On appelle **bruit blanc** (par référence à la lumière blanche) un signal chaotique dont le spectre de puissance est constant, au moins dans une très large bande de fréquences ; dans ce cas $\tau_c \simeq 0$.

EXEMPLE. En optique une **lumière quasi-monochromatique** correspond à

$$I(\nu) = \widehat{P}(\nu - \nu_0) > 0\ ,$$

où le **profil de la raie** $\widehat{P}$ a une largeur $\Delta\nu \simeq \frac{1}{\tau_c} \ll \nu_0$. On a (théorème ci-dessus)

$$\Gamma(\tau) = \Re e\left[\int_0^{\infty} \widehat{P}(\nu - \nu_0)\, e^{i2\pi\nu\tau}\,\mathrm{d}\nu\right] \simeq \Re e\left[e^{i2\pi\nu_0\tau}\, P(\tau)\right]$$

où $P(\tau) = \int_{-\infty}^{\infty} \widehat{P}(\nu')\, e^{i2\pi\nu'\tau}\,\mathrm{d}\nu'$ est une fonction lentement variable de τ (par exemple $P(\tau) = \Delta\nu\,\mathrm{sinc}(\pi\,\Delta\nu\,\tau)$ pour $\widehat{P}(\nu - \nu_0) = \Pi\left(\frac{\nu - \nu_0}{\Delta\nu}\right)$). Dans une expérience d'interférence, où le détecteur mesure $< |f(t) + f(t+\tau)|^2 >= 2\Gamma(0) + 2\Gamma(\tau)$ (*cf.* section 2.5.1.), le caractère quasi-monochromatique du signal se traduit par $\Gamma(\tau) \simeq \Gamma(0)\cos 2\pi\nu_0\tau$ pour $\tau \ll \tau_c$ (contraste des franges proche de 1), tandis que le caractère chaotique se traduit par

$$\frac{\Gamma(\tau)}{\Gamma(0)} \simeq \frac{|P(\tau)|}{P(0)} \simeq 0$$

pour $\tau \gg \tau_c$ (contraste nul aux grands ordres $\nu_0\tau$ d'interférence).

Remarques. 1) L'étude de $P(\tau)$ renseigne sur le profil de la raie : il est gaussien si $P(\tau)$ est gaussien, lorentzien si $P(\tau) \propto e^{-\frac{|\tau|}{\tau_c}}$, symétrique si $P(\tau)$ est réel, *etc.* (Voir aussi la section 10.3.3. pour le profil lorentzien.) 2) Si on modélise le signal lumineux par une somme de **paquets d'ondes** (réels) émis à des instants aléatoires $f(t) = \sum_i \varphi(t - t_i)$ on peut montrer (à partir de $\hat{f}_{T_0}(\nu) = \widehat{\varphi}(\nu)\sum_i e^{-i2\pi\nu t_i}$ avec $t_i \in [-\frac{T_0}{2}, \frac{T_0}{2}]$) que $I(\nu) = n\,2|\widehat{\varphi}(\nu)|^2$ où n est le nombre moyen de paquets par unité de temps.

▷ Filtrage des signaux chaotiques

Il découle du théorème ci-dessus et de la relation entrée-sortie $\hat{s} = \widehat{R}\,\hat{e}$ que la relation entre les spectres de puissance des signaux chaotiques à l'entrée et à la sortie d'un filtre est :

$$I_s(\nu) = |\widehat{R}(\nu)|^2\, I_e(\nu)\ .$$

Si $e(t)$ est un bruit blanc ($I_e(\nu) \simeq 1$, $\Gamma_e(\tau) \simeq \delta(\tau)$) on a pour la sortie $I_s(\nu) \simeq |\widehat{R}(\nu)|^2$ donc $\Gamma_s(\tau) \simeq \Gamma_R(\tau)$, ce qui permet un accès aux propriétés du filtre (moins complet mais aussi moins "brutal" qu'avec $\delta(t)$).

EXEMPLE 1. **Filtrage passe bas**. L'**équation de Langevin** $m\dot{\vec{v}}(t) + \alpha\vec{v}(t) = \vec{f}(t)$ (bruit blanc) décrit le **mouvement brownien** d'une particule soumise aux chocs moléculaires. Pour "le signal de sortie" $\vec{v}(t)$, on a

$$I_v(\nu) \propto |\hat{R}(\nu)|^2 \propto (1 + 4\pi^2\nu^2\tau_c^2)^{-1} \quad \text{avec} \quad \tau_c = \tfrac{m}{\alpha}$$

et donc $\Gamma_v(\tau) \propto e^{-\frac{|t|}{\tau_c}}$; le coefficient de proportionnalité est $\Gamma_v(0) =< v^2 >$. Le mouvement peut donc être assimilé à une marche au hasard (*cf.* section 10.3.1) constituée de pas de longueur $l \simeq \tau_c\sqrt{< v^2 >}$ séparés par un intervalle de temps $T \simeq \tau_c$; on obtient pour le coefficient de diffusion $D \simeq < v^2 > \tau_c \simeq \frac{k_B T}{\alpha}$ (**relation d'Einstein**).

EXEMPLE 2. **Filtrage passe bande**. Dans les **télécommunications** on sait que la modulation de fréquence $x(t) \to A\ \cos 2\pi(\nu_0 t + \Delta\nu \int^t x(s)\,\mathrm{d}s)$ est moins sensible au bruit que la modulation d'amplitude $x(t) \to A\ (1 + mx(t))\ \cos 2\pi\nu_0 t$. Pour le comprendre on remarque que, puisque lors de la détection on ne filtre que la bande $[\nu_0 - B, \nu_0 + B]$, où $[-B, B]$ est le spectre de $x(t)$, il suffit d'étudier l'effet sur $A\cos 2\pi\nu_0 t$ de l'addition d'un bruit de la forme $a(t)\cos 2\pi\nu_0 t + b(t)\sin 2\pi\nu_0 t$ où les spectres de $a(t)$ et $b(t)$ sont dans $[-B, B]$. Sous l'hypothèse que le bruit est faible, l'addition se traduit par :

$$A\cos 2\pi\nu_0 t \longrightarrow A\big(1 + \tfrac{a(t)}{A}\big)\ \cos\big(2\pi\nu_0 t - \tfrac{b(t)}{A}\big)\ .$$

En modulation d'amplitude il faut donc comparer $m^2 < x^2 >$ à $\frac{<a^2>}{A^2}$, et en modulation de fréquence $4\pi^2\,(\Delta\nu)^2\ < x^2 >$ à $\frac{<\dot{b}^2>}{A^2} = \frac{1}{A^2}\int_{-B}^{B}(2\pi\nu)^2\,|\hat{b}(\nu)|^2\,\mathrm{d}\nu\ <\ 4\pi^2B^2\,\frac{<b^2>}{A^2}$. Toutes choses égales par ailleurs, on a m^2 d'un côté et $\big(\frac{\Delta\nu}{B}\big)^2$ de l'autre ; typiquement $m \simeq 10^{-1}$ et $\dfrac{\Delta\nu}{B} \simeq 5$, soit une amélioration par un facteur 2500 du rapport signal sur bruit.

▷ Langage probabiliste (figure 21)

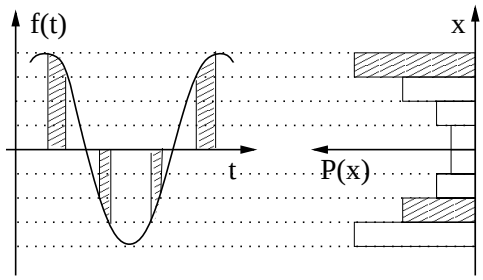


FIGURE 21

Il s'introduit naturellement quand on s'intéresse (dans l'esprit de l'intégration définie par Lebesgue) à la **fraction de temps** $P(\mathcal{D})$ pendant laquelle un signal $f(t)$ a une valeur x comprise dans un intervalle $\mathcal{D}$. ($P(\mathcal{D})$ est aussi la moyenne temporelle $< \chi_{\mathcal{D}}\,(f(t)) >$ où $\chi_{\mathcal{D}}(x) = 1$ si $x \in \mathcal{D}$ et 0 si $x \notin \mathcal{D}$.) $P(\mathcal{D})$ satisfait en effet les axiomes des probabilités, par exemple :

$$P(\mathcal{D}_1 \cup \mathcal{D}_2) = P(\mathcal{D}_1) + P(\mathcal{D}_2) \quad \text{ si } \mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset \quad , \quad P(\mathbb{R}) = 1\ .$$

Si $\mathcal{D} = \Big]x - \dfrac{\mathrm{d}x}{2}\,,\, x + \dfrac{\mathrm{d}x}{2}\Big[$ est infinitésimal, cette fraction s'écrit en général $P(x)\,\mathrm{d}x$; pour $f(t) = A\cos(\omega t + \varphi)$, on déduit de $|\mathrm{d}f| = \omega\,\sqrt{A^2 - f^2}\,|\mathrm{d}t|$ et de $P(x)\,\mathrm{d}x = 2\dfrac{|\mathrm{d}t|}{T}$

que $P(x) = \dfrac{1}{\pi\sqrt{A^2 - x^2}}$ (figure 21). Si $f(t)$ passe une fraction de temps finie p_i sur une valeur x_i, alors $P(x) = p_i\,\delta(x - x_i) + \cdots$.

A l'aide de $P(x)$ on peut traduire les **moyennes temporelles** en **moyennes probabilistes** ; en effectuant un réarrangement des intervalles de temps sur lesquels on intègre, on obtient par exemple $< f^n(t) >= \int_{-\infty}^{\infty} x^n\,P(x)\,\mathrm{d}x$ ou $< e^{iuf(t)} >= \int_{-\infty}^{\infty} e^{iux}\,P(x)\,\mathrm{d}x$ (fonction caractéristique des valeurs x du signal). De même on peut récrire la fonction de corrélation d'un signal

$$\Gamma_f(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{t_0-\frac{T}{2}}^{t_0+\frac{T}{2}} f(t)\,f(t+\tau)\,\mathrm{d}t = \iint xy\,P_\tau(x,y)\,\mathrm{d}x\,\mathrm{d}y\ ,$$

en introduisant la fraction de temps $P_\tau(x,y)\,\mathrm{d}x\,\mathrm{d}y$ pendant laquelle $f(t) \in \left]x - \dfrac{\mathrm{d}x}{2}\,,\ x + \dfrac{\mathrm{d}x}{2}\right[$ *et* $f(t+\tau) \in \left]y - \dfrac{\mathrm{d}y}{2}\,,\ y + \dfrac{\mathrm{d}y}{2}\right[$ (*cf.* densité de probabilité conjointe : section 10.2.1).

Signaux à perte de mémoire. Dans le langage probabiliste, l'indépendance des valeurs du signal à deux instants séparés de τ se traduit par la factorisation $P_\tau(x,y) = P(x)\,P(y)$, et donc pour des signaux de moyenne nulle par $\Gamma_f(\tau) = 0$. Donc la relation $\Gamma_f(\tau) \to 0$ pour $\tau \to \infty$, qui caractérise les signaux chaotiques du point de vue de l'analyse de Fourier, est une condition nécessaire (dont on se contente souvent !) pour pouvoir affirmer que ce sont des signaux "à perte de mémoire". Elle entraîne que toute moyenne temporelle pour ces signaux peut être calculée sur des durées T finies pourvu qu'elles soient grandes devant le temps de corrélation τ_c ; l'erreur relative commise est alors de l'ordre de $\dfrac{1}{\sqrt{N}}$ avec $N = \dfrac{T}{\tau_c}$ (comme pour une somme de N variables aléatoires indépendantes ; *cf.* section 10.2.1).

5.4 OPTIQUE DE FOURIER ; FILTRAGE OPTIQUE

5.4.1 Décomposition en ondes planes ; filtrage

▷ Principe général

Soit $f(x,y,0)$ l'amplitude lumineuse (monochromatique) d'un **objet cohérent** émettant vers les $z > 0$, et dont le développement de Fourier s'écrit

$$f(x,y,0) = \iint \hat{f}(\sigma_x,\sigma_y)\,e^{i2\pi(\sigma_x x + \sigma_y y)}\,\mathrm{d}\sigma_x\,\mathrm{d}\sigma_y\ ,$$

avec $\hat{f}(\sigma_x,\sigma_y) = \iint f(x,y,0)\,e^{-i2\pi(\sigma_x x + \sigma_y y)}\,\mathrm{d}x\,\mathrm{d}y$. On peut aussi l'écrire

$$f(x,y,0) = \iint \hat{f}\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\,e^{i\frac{2\pi}{\lambda}(\alpha x + \beta y)}\,\frac{\mathrm{d}\alpha}{\lambda}\,\frac{\mathrm{d}\beta}{\lambda}\ ,$$

et interpréter $f(x,y,0)$ comme représentant dans le plan $z = 0$ une **somme d'ondes planes** allant dans des directions $\hat{u}(\alpha,\beta,\gamma)$ données par $\alpha = \lambda\sigma_x$, $\beta = \lambda\sigma_y$ ($\gamma = \sqrt{1-\alpha^2-\beta^2}$).

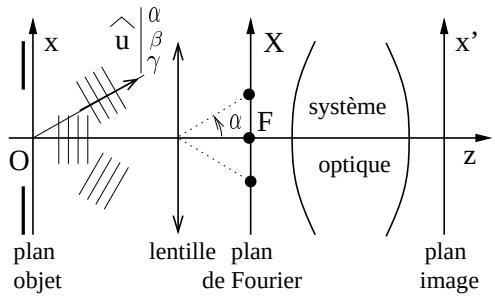


FIGURE 22

On suppose la longueur d'onde λ, ou les fréquences spatiales $\sigma_{x,y}$, assez petites pour que $\lambda\sigma_{x,y} < 1$. Si on place une lentille de distance focale d après l'objet, chaque onde plane issue de l'objet va converger dans le plan focal image $(F; X, Y)$ au point $(X = \alpha d, Y = \beta d)$ (figure 22). Cette "séparation point par point" des ondes permet un filtrage optique immédiat en mettant dans ce plan (plan de Fourier) un filtre (**pupille**) de coefficient de transmission $p(X, Y)$. Si on place alors après ce plan un système optique qui fait, avec la lentille, l'image du plan $z = 0$, cette image est celle que l'on obtiendrait (sans filtre) d'un objet (fictif) appelé **objet filtré** dont l'amplitude lumineuse serait :

$$f_{\text{filtrée}}(x, y, 0) = \iint \hat{f}(\sigma_x, \sigma_y)\, p(\lambda\sigma_x d, \lambda\sigma_y d)\, e^{i2\pi(\sigma_x x + \sigma_y y)}\, \mathrm{d}\sigma_x\, \mathrm{d}\sigma_y \ .$$

$p(\lambda\sigma_x d, \lambda\sigma_y d)$ est la **fonction de transfert du filtre**.

EXEMPLE 1. Soit $f(x, y, 0) = \cos^2 \frac{\pi x}{a} = \frac{1}{2} + \frac{1}{4}\, e^{i2\pi \frac{x}{a}} + \frac{1}{4}\, e^{-i2\pi \frac{x}{a}}$. A un trou en F correspond l'objet filtré uniforme $f_{\text{filtrée}}(x, y, 0) = \frac{1}{2}$, et à un écran ponctuel en F l'objet filtré $f_{\text{filtrée}}(x, y, 0) = \frac{1}{2} \cos \frac{2\pi x}{a}$ (dont l'intensité $\frac{1}{4} \cos^2 \frac{2\pi x}{a}$ a une période spatiale moitié de celle de l'objet initial).

EXEMPLE 2. Soit $f(x, y, 0) = A\, e^{i\varphi(x,y)} \simeq A\, e^{i\varphi_0}\, (1 + i\delta\varphi(x, y))$ un objet de phase dont on suppose les fluctuations de phase petites. Si on place en F une lame introduisant un retard de $\pm\frac{\pi}{2}$ ou un écran, l'objet filtré est $f_{\text{filtrée}}(x, y, 0) = iA\, e^{i\varphi_0}\, (\pm 1 \text{ ou } 0 + \delta\varphi(x, y))$. Il présente des variations d'intensité et son image devient observable (principe du **contraste de phase** ou de la **strioscopie**).

▷ Lien avec la diffraction de Fraunhofer

L'intensité observée dans le plan de convergence des ondes planes $(F; X, Y)$ (avant filtrage) est proportionnelle à $\left|\hat{f}\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right)\right|^2$. On retrouve l'intensité de la figure de diffraction à "l'infini" associée à $f(x, y, 0)$ et calculée autrement à la section 2.5.2.

▷ Lien avec la diffraction de Fresnel

Comme la propagation du plan $z = 0$ au plan $z > 0$ se traduit simplement pour l'onde $e^{i\frac{2\pi}{\lambda}(\alpha x + \beta y)}$ dans le plan $z = 0$ par la multiplication par le facteur $e^{i\frac{2\pi}{\lambda}\gamma z}$ $(\gamma(\alpha, \beta) = \sqrt{1 - \alpha^2 - \beta^2})$, l'amplitude $f(X, Y, z)$ dans le plan z est :

$$f(X, Y, z) = \iint \hat{f}\left(\frac{\alpha}{\lambda}, \frac{\beta}{\lambda}\right)\, e^{i\frac{2\pi}{\lambda}(\alpha X + \beta Y + \gamma(\alpha,\beta) z)}\, \frac{\mathrm{d}\alpha}{\lambda}\, \frac{\mathrm{d}\beta}{\lambda} \ .$$

Le facteur de phase $e^{i\frac{2\pi}{\lambda}\gamma z}$ (analogue de $\widehat{R}(\nu)$ pour des signaux)), qui s'écrit

$$e^{i\frac{2\pi}{\lambda}z}\, e^{-i\pi\lambda z(\sigma_x^2+\sigma_y^2)} = \widehat{R}_z(\sigma_x,\sigma_y)$$

dans l'approximation de Gauss ($\gamma \simeq 1 - \dfrac{\alpha^2+\beta^2}{2}$), n'est autre que la transformée de Fourier à deux dimensions de la réponse impulsionnelle $R_z(x,y)$ introduite à la section 5.2.3. Le **principe d'Huygens Fresnel** est donc une conséquence de la propagation des ondes planes (et de l'analyse de Fourier) et n'est pas un principe physique nouveau.

EXEMPLE. Si on éclaire avec un laser sous incidence normale un réseau périodique situé dans le plan $z=0$ de coefficient de transmission $t(x,y) = \sum_n t_n\, e^{i2\pi n\frac{x}{a}}$ $\big(= f(x,y,0)\big)$, on a $f(x,y,z) = e^{i2\pi\frac{z}{\lambda}} \times \sum_n e^{-i\pi\lambda z(\frac{n}{a})^2}\, e^{i2\pi n\frac{x}{a}}$. Dans les plans $z = 2p\frac{a^2}{\lambda}$ (p entier) on retrouve (en intensité) le réseau : $|f(x,y,2p\frac{a^2}{\lambda})| = |f(x,y,0)|$.

5.4.2 Illustrations optiques de la transformée de Fourier

Soit $t(x,y)$ le coefficient de transmission d'un écran éclairé sous incidence normale par une onde plane d'amplitude 1. On considère la relation

$$t(x,y) = \iint \hat{t}\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\, e^{i\frac{2\pi}{\lambda}(\alpha x+\beta y)}\, \frac{\mathrm{d}\alpha}{\lambda}\,\frac{\mathrm{d}\beta}{\lambda}$$

qui exprime la décomposition en ondes planes de l'amplitude transmise. Le lecteur est encouragé à représenter les "écrans" t et les "intensités" $|\hat{t}|^2$ dans les plans (x,y) et $(\frac{\alpha}{\lambda},\frac{\beta}{\lambda})$ pour les exemples ci-dessous (*cf.* ouvrages d'optique).

▷ Intensité diffractée ; ouverture rectangulaire

La **formule de Parseval** $\displaystyle\iint |t(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y = \iint \left|\hat{t}\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\right|^2 \frac{\mathrm{d}\alpha}{\lambda}\,\frac{\mathrm{d}\beta}{\lambda}$ montre que

$$I(\alpha,\beta) = \frac{1}{\lambda^2}\left|\hat{t}\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\right)\right|^2$$

est l'intensité lumineuse diffractée (à l'infini) dans la direction (α,β) par unité d'angle solide. Un exemple classique est $t(x,y) = \Pi\left(\dfrac{x}{a}\right)\,\Pi\left(\dfrac{y}{b}\right)$ (ouverture rectangulaire), qui conduit à

$$I(\alpha,\beta) = \left(\frac{ab}{\lambda}\right)^2 \left(\mathrm{sinc}\frac{\pi a\alpha}{\lambda}\right)^2 \left(\mathrm{sinc}\frac{\pi b\beta}{\lambda}\right)^2$$

(l'intégrale double étant ici un produit d'intégrales simples).

▷ Linéarité ; réalité ; aires

Deux écrans sont dits "complémentaires" si $t_1 + t_2 = 1$ (fente fine et fil opaque) ; alors $\hat{t_1} = \delta - \hat{t_2}$ et les intensités diffractées sont les mêmes (sauf au centre $\alpha=\beta=0$) : **théorème de Babinet**.

Si $t(x,y)$ est réel (**écran réel**), $I(\alpha,\beta)=I(-\alpha,-\beta)$: la figure de diffraction est symétrique même si l'écran ne l'est pas ! Si $t(x,y)>0$ alors $I(\alpha,\beta)$ est maximum au centre ($\alpha=\beta=0$) ; en particulier si l'écran est un trou de surface S, alors $I(0,0)=\lambda^{-2}\left|\iint t(x,y)\,\mathrm{d}x\,\mathrm{d}y\right|^2 = S^2/\lambda^2$, tandis que l'intensité totale, intégrée sur toutes les directions vaut S ; ceci se comprend en remarquant que la tache de diffraction a une extension angulaire $\Delta\alpha\,\Delta\beta \simeq \lambda^2/S$.

▷ Dilatation

Un changement d'échelles sur l'écran $t(x,y) \to t\Big(\dfrac{x}{k_1},\dfrac{y}{k_2}\Big)$ se traduit par $I(\alpha,\beta) \to k_1^2\,k_2^2\,I(k_1\alpha,k_2\beta)$ (changement d'échelles inverse) ; en particulier si $k_2 \to \infty$ (ouverture très large en y), $I(\alpha,\beta)$ se concentre sur les directions $\beta=0$.

▷ Translation de l'écran

$\hat{t}$ est multiplié par une phase et la figure de diffraction n'est donc pas modifiée. Par contre le facteur de phase se manifeste lors de la diffraction par plusieurs écrans translatés les uns par rapport aux autres. Par exemple à $\sum_{n=1}^{N} t(x-na,y)$ correspond $I(\alpha,\beta)\,|\sum_{n=1}^{N} e^{in\varphi}|^2$ produit de $I(\alpha,\beta)$ par l'intensité d'une interférences à N trous ponctuels ($\varphi = -\dfrac{2\pi\alpha a}{\lambda}$).

▷ Multiplication de $t(x,y)$ par $e^{i\frac{2\pi}{\lambda}(\alpha_0 x+\beta_0 y)}$

Elle correspond à un **éclairage sous incidence oblique** et conduit à $I(\alpha-\alpha_0,\beta-\beta_0)$ (décalage angulaire de la figure de diffraction).

▷ Périodisation

Un **réseau infini** a un facteur de transmission périodique

$$t_\infty(x) = \sum_{n=-\infty}^{\infty} \tau(x-na) = \frac{1}{a}\sum_{p=-\infty}^{\infty} \hat{\tau}\big(\tfrac{p}{a}\big)\,e^{i2\pi p\frac{x}{a}}$$

(formule de Poisson). $\tau(x)$ (restreint à $[0,a]$) caractérise le **motif du réseau** ; l'enjeu des **réseaux "blazés"** est de choisir judicieusement le motif $\tau(x)$ pour concentrer toute la lumière diffractée dans un seul ordre q (pour augmenter la luminosité), le plus élevé possible pour avoir le pouvoir séparateur maximum : $\hat{\tau}\left(\dfrac{p}{a}\right)=0$ sauf pour $p=q$. Un exemple est $\tau(x)=e^{i2\pi\frac{x}{a}}$. L'écran correspondant est un prisme d'angle au sommet θ petit et d'indice n tel que $\dfrac{2\pi}{\lambda}\,(n-1)\,\theta = 2\pi\,\dfrac{q}{a}$.

REMARQUE. Si le réseau est limité par une ouverture de coefficient de transmission $\mathcal{O}(x,y)$, alors :

$$t(x,y)=t_\infty(x)\,\mathcal{O}(x,y) \text{ et } \hat{t}\big(\tfrac{\alpha}{\lambda},\tfrac{\beta}{\lambda}\big)=\tfrac{1}{a}\sum_{p=-\infty}^{\infty}\hat{\tau}\big(\tfrac{p}{a}\big)\,\widehat{\mathcal{O}}\big(\tfrac{\alpha-\alpha_p}{\lambda},\tfrac{\beta}{\lambda}\big)\ .$$

Cette formule peut s'interpréter comme la diffraction des différentes ondes planes $e^{i2\pi\alpha_p\frac{x}{\lambda}}$ par l'ouverture. Si celle-ci est plus large que le **pas du réseau** a, les fonctions $\widehat{\mathcal{O}}\big(\frac{\alpha-\alpha_p}{\lambda},\frac{\beta}{\lambda}\big)$ (pour les différentes valeurs de p) ne se recouvrent pas et l'intensité diffractée est $\frac{1}{a^2}\sum_{p=-\infty}^{\infty}\big|\hat{\tau}\big(\frac{p}{a}\big)\big|^2\,I_{\mathcal{O}}(\alpha-\alpha_p,\beta)$, où $I_{\mathcal{O}}(\alpha,\beta)$ désigne l'intensité diffractée par l'ouverture.

▷ Echantillonnage

Si $t_\infty(x)$ correspond à un réseau de fentes fines ($\hat\tau\big(\frac{p}{a}\big)$ =constante), on peut considérer que $t_\infty(x)\,\mathcal{O}(x,y)$ est l'échantillonnée en x de "l'objet "$\mathcal{O}(x,y)$; $\hat t\big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\big)$ est alors la périodisée de $\widehat{\mathcal{O}}\big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\big)$. Dans le cas de non recouvrement (condition de Shannon satisfaite), un filtrage passe-bas ne laissant passer que le terme $\widehat{\mathcal{O}}\big(\frac{\alpha}{\lambda},\frac{\beta}{\lambda}\big)$ permet de récupérer $\mathcal{O}(x,y)$ (principe du **détramage**).

▷ Discontinuité

Si $t(x,y)$ est discontinu le long d'une ligne $x=0$ la figure de diffraction s'étend dans la direction α sur l'axe $\beta=0$. La technique d'**apodisation** consiste à réduire cette génération de hautes fréquences, par exemple en remplaçant $t(x)=\Pi\big(\frac{x}{a}\big)$ par $t_{\rm apod}(x)=\cos^2\frac{\pi x}{a}\,\Pi\big(\frac{x}{a}\big)$, dont la T.F. (petit calcul) $\widehat{t_{\rm apod}}(\sigma)=\frac{a}{2(1-a^2\sigma^2)}\,{\rm sinc}\pi\sigma a$ est un peu plus étalée que $\hat t(\sigma)$ et décroît comme σ^{-3} au lieu de σ^{-1} (discontinuité de $t''_{\rm apod}$).

▷ Autocorrélation

EXEMPLE 1 : **optique incohérente**. Les exemples ci-dessus relevaient de l'optique cohérente. Si une répartition angulaire d'intensité $I(\alpha_0,\beta_0)$ en provenance d'un objet incohérent à l'infini arrive dans la pupille d'entrée d'un instrument de surface S et de coefficient de transmission $t(x,y)$, l'intensité à la sortie de la pupille est donnée par le produit convolution (normalisé de sorte que $I_f=I$ si I est uniforme et $t=1$)

$$I_f(\alpha,\beta)=\iint\ \tfrac{1}{S\lambda^2}\,I(\alpha_0,\beta_0)\,\big|\hat t\big(\tfrac{\alpha-\alpha_0}{\lambda},\tfrac{\beta-\beta_0}{\lambda}\big)\big|^2\,{\rm d}\alpha_0\,{\rm d}\beta_0\ .$$

(I_f résulte de l'addition des intensités des figures de diffraction associées chacune à une onde plane). La pupille effectue donc un filtrage sur l'objet. Si on introduit les T.F. des intensités I et I_f par

$$\widehat I(\sigma_\alpha,\sigma_\beta)=\iint\ I(\alpha,\beta)\,e^{-i2\pi(\sigma_\alpha\alpha+\sigma_\beta\beta)}\,{\rm d}\alpha\,{\rm d}\beta$$

(σ_α et σ_β fréquences angulaires), on a $\widehat{I_f}=T\widehat I$ avec pour **fonction de transfert de l'instrument**

$$T(\sigma_\alpha,\sigma_\beta)=\tfrac{1}{S}\iint\,\overline{t}(x,y)\,t(x-\lambda\sigma_\alpha,y-\lambda\sigma_\beta)\,{\rm d}x\,{\rm d}y$$

la fonction d'**autocorrélation de la pupille**. Comme une pupille a toujours une extension finie, $T(\sigma_\alpha,\sigma_\beta)$ a un support borné et l'instrument coupe les hautes fréquences. Si il s'agit de deux trous distants de a dans la direction x, il ne laisse passer que les fréquences $\sigma_\alpha=\sigma_\beta=0$ et $\sigma_\alpha=\pm\dfrac{a}{\lambda}$, $\sigma_\beta=0$. On peut ainsi étudier $\hat I(\sigma_\alpha,\sigma_\beta)$ et remonter à $I(\alpha,\beta)$ (principe de l'**interférométrie stellaire**).

EXEMPLE 2 : **diffusion**. Si N diffuseurs isotropes identiques, occupant les positions $\overrightarrow{r}_i$ dans un volume, sont éclairés par une onde plane $e^{i\overrightarrow{k}_0\cdot\overrightarrow{r}}$, l'intensité diffusée dans la direction du vecteur d'onde $\overrightarrow{k}$ est proportionnelle à :

$$S(\overrightarrow{\sigma})=<|\Sigma_i\,e^{-i2\pi\overrightarrow{\sigma}\cdot\overrightarrow{r}_i}|^2>\qquad(2\pi\overrightarrow{\sigma}=\overrightarrow{k}-\overrightarrow{k}_0)\ .$$

(La moyenne $<\ \cdot\ >$ est à prendre si les positions des diffuseurs sont aléatoires.) Comme $\Sigma_i\,e^{-i2\pi\overrightarrow{\sigma}\cdot\overrightarrow{r}_i}$ est la T.F. (tridimensionnelle) de la densité de diffuseurs

$$n(\overrightarrow{r})=\Sigma_i\,\delta(\overrightarrow{r}-\overrightarrow{r}_i)\ ,$$

$S(\overrightarrow{\sigma})$ est la T.F. de la fonction d'autocorrélation de la densité :

$$<\int\,n(\overrightarrow{r}')\,n(\overrightarrow{r}'+\overrightarrow{r})\,{\rm d}^3r'>=<\textstyle\sum_{ij}\delta(\overrightarrow{r}-\overrightarrow{r}_i+\overrightarrow{r}_j)>\simeq\ N\,g(\overrightarrow{r})\ .$$

$g(\overrightarrow{r})\,{\rm d}^3r$ est le nombre moyen de diffuseurs dans ${\rm d}^3r$, sachant qu'il y en a un en $\overrightarrow{r}=0$.

Chapitre 6

Equations différentielles ; systèmes dynamiques

Une **équation différentielle (E.D.)** ordinaire d'ordre n se présente comme une relation, en général non linéaire, entre une fonction d'une variable et ses dérivées d'ordre $1, \cdots, n$. La résoudre consiste traditionnellement à en donner la solution générale, mais ceci est rarement réalisable à part quelques cas classiques, et fait perdre de vue l'information "codée" par les E.D. en physique. C'est pourquoi il est utile de commencer par l'approche qualitative et géométrique des systèmes dynamiques (où la variable est le temps). Cette approche repose sur la propriété d'un système d'E.D. d'ordres quelconques d'être équivalent à une E.D. "vectorielle" du premier ordre décrivant le mouvement d'un point dans l'espace de phase. Elle met en valeur le rôle important des **conditions initiales (C.I.)**. On verra que, lorsque la variable est l'espace, ce rôle est tenu plutôt par les **conditions aux limites (C.L.)**.

Quand une solution évidente est connue (par exemple l'état de repos pour un pendule), il est naturel de s'intéresser à l' E.D. satisfaite au premier ordre par une perturbation (petites oscillations). C'est pourquoi les **E.D. linéaires (E.D.L.)** jouent un rôle privilégié. On étudiera leurs propriétés, non seulement lorsque les coefficients sont constants, cas des **E.D.L. stationnaires (E.D.L.S)**, avec la notion de mode propre stable ou instable, mais aussi dans le cas général, avec la notion de matrice de transfert. L'approche linéaire ayant ses limites, on étudiera enfin quelques méthodes, notamment la méthode des équations d'amplitude, permettant de traiter les corrections non linéaires et les bifurcations de solutions.

6.1 SYSTÈMES DYNAMIQUES ET ESPACE DE PHASE

6.1.1 Définitions ; propriétés générales

▷ Systèmes dynamiques

L'idée de **déterminisme**, dans sa version "continue", se traduit par le fait que la connaissance de l'**état** $\underline{x}(t_0)$ d'un système à un instant arbitraire t_0 (conditions initiales) entraîne sa connaissance à un instant voisin $t_0 + \epsilon$, ce qu'on écrit :

$$\underline{x}(t_0+\epsilon) \simeq \underline{x}(t_0) + \epsilon \underline{f}\big(\underline{x}(t_0), t_0\big) \ \forall t_0 \text{ et } \underline{x}(t_0) \qquad \text{soit lorsque } \epsilon \to 0 \qquad \dot{\underline{x}}(t) = \underline{f}\big(\underline{x}(t), t\big) \ .$$

L'état est ainsi connu de proche en proche, au moins dans un certain intervalle de temps, aussi bien dans le futur que dans le passé, dès que sa "vitesse d'évolution" $\underline{f}(\underline{x}, t)$ l'est. On parle alors de **système dynamique (S.D.)** dont l'**évolution** est régie par l'E.D. "vectorielle" du premier ordre.

Attention. Une E.D. décrit l'évolution, non d'un état initial particulier, mais celle de n'importe quel état initial $\underline{x}(t_0)$. Elle "code" une correspondance générale, souvent très compliquée, entre états pour deux instants différents (figure 1) :

$$\underline{x}(t_0) \xrightarrow{F_{t,t_0}} \underline{x}(t) \ .$$

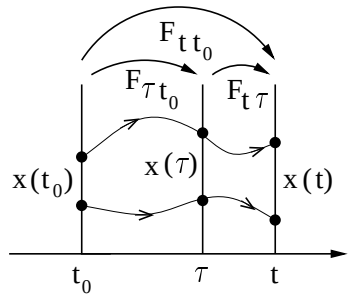


FIGURE 1

Cette correspondance n'est pas quelconque ; elle doit satisfaire la loi de composition :

$$F_{t,t_0} = F_{t,\tau} \ \circ \ F_{\tau,t_0} \ .$$

Par exemple la relation $x(t) = a\,x(t_0) + (t - t_0)$ ne satisfait cette loi que si $a = 1$. En effet, en prenant un instant quelconque τ, on doit avoir $x(t) = a\,x(\tau) + (t - \tau) = a\big(a\,x(t_0) + (\tau - t_0)\big) + (t - \tau)$. Le cas $a = 1$ correspond au S.D. $\dot{x} = 1$.

▷ Espace de phase ; degrés de liberté

La notation $\underline{x}$ sous entend que pour caractériser l'état d'un système il faut, en général, se donner plusieurs grandeurs (ses composantes $x_1, x_2, \cdots, x_n$). En mécanique il faut connaître initialement les positions et les vitesses des corps, en électricité les charges des

condensateurs et les intensités dans les inductances, en cinétique chimique toutes les concentrations, *etc.* Ces grandeurs (fonctions du temps) qui caractérisent un état définissent l'espace de phase, espace de tous les états possibles, dans lequel tout état initial décrira une "trajectoire" au cours du temps. La dimension de cet espace caractérise le nombre de degrés de liberté du S.D.

L'intérêt de la notion d'état dans l'espace de phase est de ramener toute E.D. (ou système d'E.D.) de n'importe quel ordre à un **système du premier ordre** (pour les composantes de l'état). Par exemple l'équation du troisième ordre $\dddot{x} = g(x, \dot{x}, \ddot{x}, t)$ s'écrit aussi

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ g(x_1, x_2, x_3, t) \end{pmatrix} \quad \Big(\text{de la forme } \dot{\underline{x}}(t) = \underline{f}\big(\underline{x}(t), t\big)\Big) \ ,$$

où $x_1 = x,\ x_2 = \dot{x},\ x_3 = \ddot{x}$ sont les composantes de l'état $\underline{x}$, et x_2, x_3, g celles de $\underline{f}$; il y a trois degrés de liberté et il faut trois C.I. $x_1(t_0) = x(t_0),\ x_2(t_0) = \dot{x}(t_0),\ x_3(t_0) = \ddot{x}(t_0)$ pour définir une solution $x(t)$.

Inversement il est à noter qu'en physique toutes les lois "de base" ne font intervenir que les dérivées premières des grandeurs : $I = \dot{Q}$, $V = L\dot{I}$ en électricité, $\vec{p} = m\dot{\vec{r}}$, $\dot{\vec{p}} = \overrightarrow{F}(\vec{r}, \vec{v}, t)$ en mécanique (loi de Newton), *etc.* Ce n'est que parce qu'on procède habituellement par "élimination" pour résoudre ces systèmes d'E.D. du premier ordre, qu'apparaissent des E.D. d'ordre supérieur, dont la difficulté technique fait du coup perdre de vue le rôle des C.I.

▷ Stationnarité

Lorsque t n'apparaît pas explicitement dans $\underline{f}(\underline{x}, t)$, c'est-à-dire lorsque les coefficients des E.D. (qui sont les paramètres "caractéristiques" du système physique, la masse d'un corps ou la raideur d'un ressort par exemple) ne dépendent pas du temps, on parle d'un **S.D. stationnaire (S.D.S.)**. Dans ce cas, la relation entre les états $\underline{x}(t_0)$ et $\underline{x}(t)$ ne fait intervenir les instants t_0 et t qu'à travers leur différence $t - t_0$ car il n'y a pas d'origine des temps privilégiée.

EXEMPLE : $\dot{x} = x(1 - x)$, qui décrit l'évolution d'une "population" x (biologique, chimique, *etc.*) dont le taux de variation $1 - x$ décroit quand x croit. L'intégration de

$$\mathrm{d}t = \big(x(1-x)\big)^{-1}\,\mathrm{d}x = \big(x^{-1} + (1-x)^{-1}\big)\,\mathrm{d}x = \mathrm{d}\big(\ln x(1-x)^{-1}\big)$$

donne $x(t) = x(t_0)\exp(t - t_0)\,\big(1 + x(t_0)\big(\exp(t - t_0) - 1\big)\big)^{-1}$.

CONTRE EXEMPLE : $\dot{x} = \lambda(t)\,x$, où le taux de variation $\lambda(t)$ de la "population" x dépend explicitement du temps. Alors $x(t) = x(t_0)\exp\int_{t_0}^{t} \lambda(s)\,\mathrm{d}s$; on ne retrouve la dépendance en $t - t_0$ que si λ est constant.

D'un point de vue géométrique $\underline{f}(\underline{x}, t)$ représente un **champ de vecteurs** dans l'espace de phase. Dans le cas stationnaire, les trajectoires suivies par les états sont les **lignes de champ**, lignes partout tangentes aux vecteurs $\underline{f}(\underline{x})$, puisque $\dot{\underline{x}} = \underline{f}(\underline{x})$. Le tableau ci-dessous et les figures 2 présentent quelques exemples d'équations adimensionnées tirées de la mécanique ; les variables de l'espace de phase sont $(x, \dot{x})$. La constante C est une intégrale première des équations.

S.D.	(a) mouvement inertiel	(b) mouvement unif. accéléré	(c) pendule près du pt. d'équilibre stable	(d) pendule près du pt. d'équilibre instable
E.D.	$\ddot{x} = 0$	$\ddot{x} = 1$	$\ddot{x} = -x$	$\ddot{x} = x$
champ de vecteurs $\underline{f}$	$\begin{pmatrix} \dot{x} \\ 0 \end{pmatrix}$	$\begin{pmatrix} \dot{x} \\ 1 \end{pmatrix}$	$\begin{pmatrix} \dot{x} \\ -x \end{pmatrix}$	$\begin{pmatrix} \dot{x} \\ x \end{pmatrix}$
trajectoires dans l'espace de phase	$\dot{x} = C$ droites	$x = \frac{1}{2}\dot{x}^2 + C$ paraboles	$x^2 + \dot{x}^2 = C$ cercles	$x^2 - \dot{x}^2 = C$ hyperboles
x et $\dot{x}$ fonction des C.I. à $t = 0$	$x = x_0 + \dot{x}_0 t$ $\dot{x} = \dot{x}_0$	$x = x_0 + \dot{x}_0 t + \frac{1}{2}t^2$ $\dot{x} = \dot{x}_0 + t$	$x = x_0 \cos t + \dot{x}_0 \sin t$ $\dot{x} = \dot{x}_0 \cos t - x_0 \sin t$	$x = x_0 \cosh t + \dot{x}_0 \sinh t$ $\dot{x} = \dot{x}_0 \cosh t + x_0 \sinh t$
représentation matricielle de $F_{t,0}$	$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$		$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$	$\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$

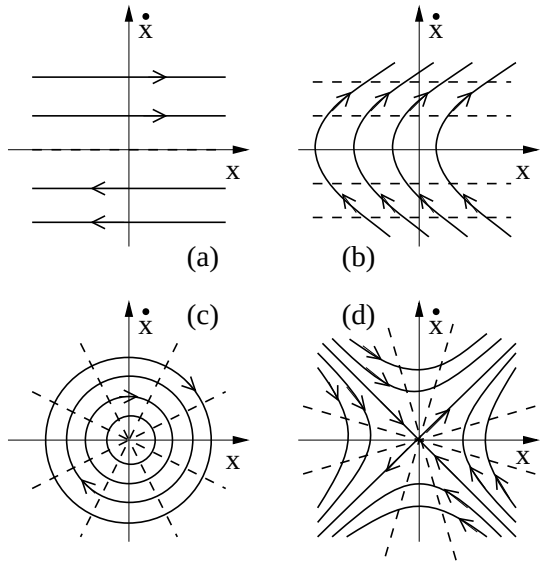


FIGURE 2

▷ Linéarité

Un S.D. linéaire (S.D.L.) est un système tel que si $\underline{x}_1(t)$ et $\underline{x}_2(t)$ sont deux solutions arbitraires $\lambda_1 \underline{x}_1(t) + \lambda_2 \underline{x}_2(t)$ est aussi solution. L'espace de phase a alors une structure d'**espace vectoriel**.

EXEMPLE : $\ddot{x} + a(t)\,\dot{x} + b(t)\,x = 0 \iff \dot{\underline{x}} = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \underline{x}$ avec $\underline{x} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$.

CONTRE EXEMPLE : $\ddot{x} = 1$ (*cf.* tableau).

La propriété essentielle d'un S.D.L. est que sa **solution générale** est une combinaison linéaire de solutions linéairement indépendantes où, ce qui est équivalent, que chaque

composante de l'état $\underline{x}(t)$ dépend linéairement des composantes de l'état $\underline{x}(t_0)$. Ainsi la solution générale de $\ddot{x} + a(t)\,\dot{x} + b(t)\,x = 0$ s'écrit

$$x(t) = R_0(t, t_0)\,x(t_0) + R_1(t, t_0)\,\dot{x}(t_0)\ ,$$

$\dot{x}(t)$ se déduisant simplement par dérivation de cette relation. Dans cette expression $R_0(t, t_0)$ est la solution qui vaut 1 pour $t = t_0$ et dont la dérivée s'annule pour $t = t_0$, tandis que $R_1(t, t_0)$ est la solution qui s'annule en $t = t_0$ et a une dérivée égale à 1 pour $t = t_0$. Cette remarque élémentaire est à la base du formalisme de la matrice de transfert développé à la section 6.4.1.

6.1.2 Exemples de systèmes dynamiques et de leurs portraits de phase

La représentation des trajectoires des états d'un S.D. dans l'espace de phase, appelée **portrait de phase**, présente l'intérêt de visualiser toutes les solutions d'une E.D., en fonction des C.I., même lorsqu'il n'y a pas de formule explicite ou que celle-ci est compliquée.

▷ Portrait de phase (à une dimension) de $\dot{x} = f(x)$

L'évolution de $x(t)$ est schématisée en plaçant simplement sur l'axe x les points où f s'annule, et en mettant entre ces points une flèche dans le sens d'évolution de $x(t)$, donc liée au signe de $f(x)$. La seule chose qui reste à préciser est de savoir si x s'approche où s'éloigne des points où f s'annule en un temps fini ou infini.

EXEMPLE 1. $\dot{x} = x(1 - x)$ et $x > 0$ (figure 3).
Près de $x = 1$ on a $\dot{x} \simeq 1 - x$, d'où $\mathrm{d}t \simeq (1 - x)^{-1}\,\mathrm{d}x$ et $t \simeq -\ln|1 - x|$; on retrouve que l'approche de 1 est exponentiellement lente. Pour x proche de 0, on a $\mathrm{d}t \simeq x^{-1}\,\mathrm{d}x$; il faut donc un temps $\Delta t \simeq |\ln \epsilon|$ pour s'écarter d'une distance finie d'une position initiale $x(t_0) = \epsilon$ arbitrairement proche de l'origine.

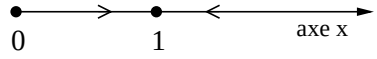


FIGURE 3

EXEMPLE 2 (figure 4a). **En mécanique**, l'équation $m\ddot{x} = -V'(x)$ ($V(x)$ énergie potentielle) dont l'espace de phase est a priori à deux dimensions, se réduit, compte tenu de la conservation de l'énergie $E = \frac{1}{2}m\dot{x}^2 + V(x)$, à $\dot{x} = \pm\sqrt{2m(E - V(x))}$. Pour $E = E_1$, x atteint ou quitte la valeur extrême x_{B_1}, telle que $V(x_{B_1}) = E_1$, en un temps fini. En effet, au voisinage de x_{B_1} l'E.D. s'écrit $\mathrm{d}t \propto \big(E_1 - V(x)\big)^{-\frac{1}{2}}\,|\mathrm{d}x| \propto |x_{B_1} - x|^{-\frac{1}{2}}\,|\mathrm{d}x|$, et donc $\Delta t \propto |x_{B_1} - x|^{\frac{1}{2}}$. La même analyse vaut pour x_{A_1} et, comme l'accélération $\ddot{x}$ ne s'annule ni en x_{B_1} ni en x_{A_1}, x oscille indéfiniment entre ces deux valeurs extrêmes.

Au contraire pour $E = E_2$, au voisinage de x_{B_2} on a $\mathrm{d}t \propto \left(E_2 - V(x)\right)^{-\frac{1}{2}} |\mathrm{d}x| \propto |x_{B_2} - x|^{-1} |\mathrm{d}x|$, et donc x_{B_2} apparaît comme une valeur limite atteinte exponentiellement lentement.

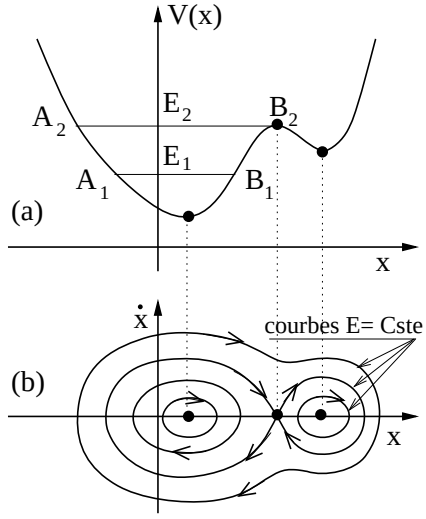


FIGURE 4

▷ Portrait de phase (à deux dimensions) de $\ddot{x} = g(x, \dot{x})$

Les remarques suivantes aident à sa construction.
a) Dans l'espace de phase paramétré par $(x, \dot{x})$ l'évolution de x (resp. $\dot{x}$) est donnée par le signe de $\dot{x}$ (resp. g). Il convient de rechercher les éventuels **points fixes** (ceux laissés invariants par l'évolution) et d'étudier leur stabilité. Par exemple pour le système mécanique ci-dessus, on a $\frac{1}{2}m\dot{x}^2 + \frac{1}{2}V''(x^*)\,(x - x^*)^2 = \mathrm{C^{ste}}$ près de toute valeur x^* telle que $V'(x^*) = 0$. Si $V''(x^*) \neq 0$, le portrait de phase (figure 4b) est localement analogue à celui des figures 2c et 2d.
b) La pente d'une trajectoire est :

$$\mathrm{tg}\,\theta = \frac{\mathrm{d}\dot{x}}{\mathrm{d}x} = \frac{\ddot{x}}{\dot{x}} = \frac{g(x, \dot{x})}{\dot{x}} \ .$$

Pour θ fixé, cette équation définit dans le plan $(x, \dot{x})$ une courbe appelée **isocline** (en pointillé sur la figure 2), le long de laquelle tous les vecteurs tangents aux différentes trajectoires sont parallèles à une direction donnée de pente $\mathrm{tg}\,\theta$. Deux trajectoires ne peuvent se couper (en vertu du déterminisme) ; elles peuvent à la rigueur tendre l'une vers l'autre (par exemple à la fin de l'évolution).

▷ Autres exemples à deux dimensions

EXEMPLE 1 : $\dot{x}_1 = k_1 x_1(1 - x_2)$, $\dot{x}_2 = -k_2 x_2(1 - x_1)$ avec k_1 et $k_2 > 0$ (**oscillateur de Lotka-Volterra d'un système proies-prédateurs**). Les proies x_1 ont un taux de

croissance $k_1(1-x_2)$ limité par les prédateurs x_2. Les prédateurs au contraire ont besoin de proies en nombre suffisant ($x_1 > 1$) pour que leur taux de croissance $-k_2(1-x_1)$ devienne positif. Si on choisit comme variables (x_1, x_2) dans le plan de phase, les points O $(0,0)$ et A $(1,1)$ sont des points fixes (figure 5). Au voisinage de O les E.D. $\dot{x}_1 \simeq k_1 x_1$ et $\dot{x}_2 \simeq -k_2 x_2$ montrent que O est **linéairement instable** : bien qu'on se rapproche de O dans la direction Ox_2 ("contractante"), on s'en éloigne dans la direction Ox_1 ("dilatante"). Au contraire le point A est **"marginalement" stable** (aucune direction dilatante ni contractante) ; dans son voisinage les E.D. s'écrivent $\dot{u}_1 \simeq -k_1 u_2$ et $\dot{u}_2 \simeq k_2 u_1$ (avec $u_{1,2} = x_{1,2} - 1$) et les trajectoires sont les ellipses centrées en A d'équations $k_2 u_1^2 + k_1 u_2^2 = \text{C}^{\text{ste}}$; leur sens de parcours est donné par le signe de $\dot{x}_1$ ou $\dot{x}_2$.

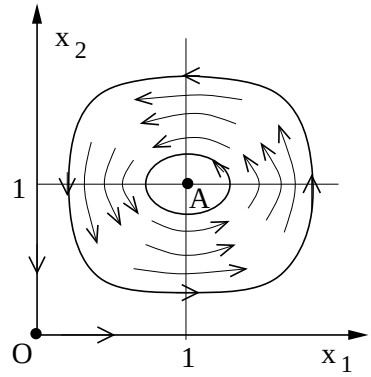


FIGURE 5

EXEMPLE 2 : $\ddot{x} = -x + F(\dot{x})$. Si la force $F(\dot{x})$ est telle que

$$F(\dot{x}) = 1 \text{ si } \dot{x} < v \quad , \quad F(\dot{x}) = -1 \text{ si } \dot{x} > v \quad , \quad -\mu \leq F(\dot{x}) \leq \mu \text{ si } \dot{x} = v \ ,$$

c'est l'équation adimensionnée d'un **oscillateur entraîné avec frottement solide** (figure 6), $\mu \geq 1$ étant le rapport des coefficients de frottement statique et dynamique.

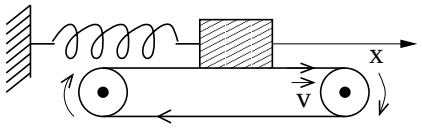
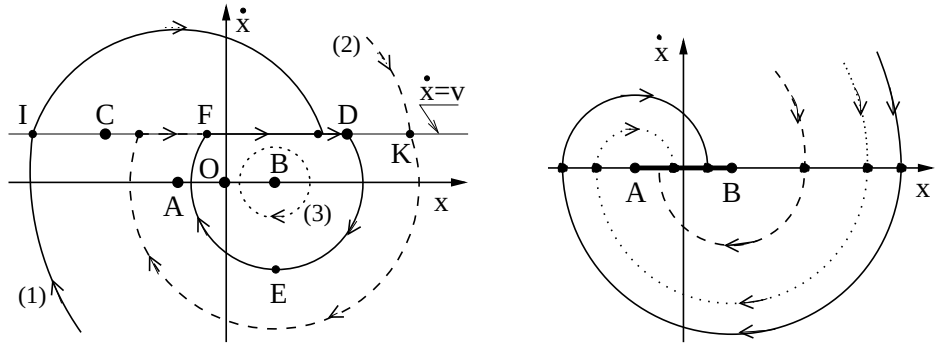


FIGURE 6

Si $\dot{x} \neq v$ alors $\ddot{x} = -x \pm 1$ et $\dot{x}^2 + (x \mp 1)^2 =$ constante. Les trajectoires sont donc des arcs de cercle centrés en A (-1,0) si $\dot{x} > v$, et des cercles (ou arcs de cercle) centrés en B (1,0) si $\dot{x} < v$ (figure 7a). Le raccord des cercles sur la ligne $\dot{x} = v$ est trivial sauf si $x \in [-\mu, \mu]$ (entre C et D). En tout point de CD le module de la force de rappel $|-x|$ est inférieur à μ. Ceci permet à la force de frottement solide $F(\dot{x})$ de "s'adapter" pour avoir $\dot{x} = v$. Ce mouvement uniforme se poursuit jusqu'en D où l'adaptation cesse d'être possible, et la trajectoire devient l'arc DEF centré en B.

A partir du point F ces trajectoires se confondent avec le **cycle** FDEF, suite d'adhérences au sol et d'oscillations liée à l'inégalité stricte $\mu > 1$. Le cas particulier $v = 0$ (pas d'entraînement) $\mu = 1$, est représenté sur la figure 7b ; les trajectoires s'arrêtent sur le segment AB.



(a) (b)

FIGURE 7

EXEMPLE 3 : $\ddot{x} + \omega^2(t)\, x = 0$ (**oscillateur paramétrique linéaire** ; figure 8). La pulsation $\omega(t)$ est périodique ($\omega(t) = \omega(t+T)$) de période $T = \dfrac{T_1 + T_2}{4}$ avec $\omega(t) = \omega_1$ pour $0 < t < \dfrac{T_1}{4}$, $\omega(t) = \omega_2$ pour $\dfrac{T_1}{4} < t < T$ et $T_{1,2} = \dfrac{2\pi}{\omega_{1,2}}$. Pour chaque ω_i la quantité $\dot{x}^2 + \omega_i^2\, x^2$ est conservée. Dans le diagramme $(\omega_0 x, \dot{x})$, où ω_0 désigne une pulsation intermédiaire ($\omega_1 > \omega_0 > \omega_2$), la trajectoire est donc une ellipse d'équation $\dot{x}^2 + \dfrac{\omega_i^2}{\omega_0^2}(\omega_0 x)^2 = \mathrm{C^{ste}}$, dont le grand axe est alternativement l'axe $\dot{x}$ (si $i = 1$) puis l'axe $\omega_0 x$ (si $i = 2$). Partant de A (1,0) à $t = 0$, on arrive donc en A_1 $(0, -\dfrac{\omega_1}{\omega_0})$ à $t = \dfrac{T_1}{4}$,

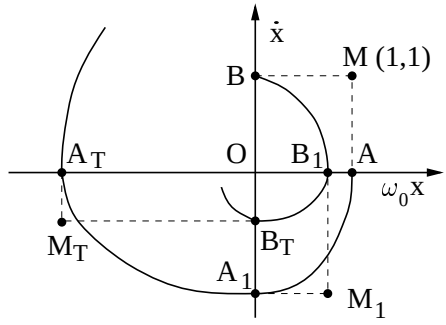


FIGURE 8

puis en A_T $(-\dfrac{\omega_1}{\omega_2}, 0)$ au bout d'une période, *etc.* (spirale divergente). Partant de B (0,1) à $t = 0$, on arrive en B_1 $(\dfrac{\omega_0}{\omega_1}, 0)$ à $t = \dfrac{T_1}{4}$, puis en B_T $(0, -\dfrac{\omega_2}{\omega_1})$ au bout d'une période,

etc. (spirale convergente). Si maintenant on part à $t = 0$ d'un point M quelconque de coordonnées $(\omega_0 x(0), \dot{x}(0))$, égales à 1 sur la figure, donc tel que $\overrightarrow{OM} = \omega_0 x(0)\,\overrightarrow{OA} + \dot{x}(0)\,\overrightarrow{OB}$, alors la linéarité de l'E.D. entraîne qu'on arrive à l'instant t en M_t tel que $\overrightarrow{OM_t} = \omega_0 x(0)\,\overrightarrow{OA_t} + \dot{x}(0)\,\overrightarrow{OB_t}$, où A_t et B_t sont les positions respectives de A et B à l'instant t. Il en résulte que, sauf si $x(0) = 0$, toutes les trajectoires divergent vers l'infini : ceci traduit le phénomène de **résonance paramétrique**.

EXEMPLE 4 : **oscillateur de Van der Pol ; oscillations de relaxation** (voir la section 6.5.2).

6.2 ÉQUATIONS LINÉAIRES STATIONNAIRES ; MODES PROPRES ; STABILITÉ

Nous commençons par rappeler les solutions bien connues des équations du premier et du second ordre. Les modes propres de toute autre équation se ramènent, en général, à l'un de ces deux cas.

6.2.1 Equations du premier et du second ordre sans et avec second membre

▷ E.D.L.S. du premier ordre réelle et complexe :

$$\dot{x} + \lambda x = 0 \quad \text{et} \quad \dot{z} + (\lambda + i\omega)z = 0 \quad (z = x + iy)\ .$$

Leurs solutions s'écrivent :

$$x(t) = x(0)\,e^{-\lambda t} \quad \text{et} \quad z(t) = z(0)\,e^{-(\lambda+i\omega)t}\ ;$$

elles correspondent respectivement, pour $\lambda > 0$, à l'exemple le plus simple de **relaxation sans oscillation** vers le point d'équilibre $x = 0$, et de **relaxation avec oscillations** (pour les variables x et y) de pulsation ω vers le point d'équilibre $x = y = 0$. Dans ce deuxième cas, l'équivalence avec l'équation habituelle d'un oscillateur amorti s'obtient en séparant parties réelle et imaginaire dans l'E.D. pour z, $\dot{x} = -\lambda x + \omega y$ et $\dot{y} = -\lambda y - \omega x$, et en éliminant une des variables (par exemple y) :

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x = 0 \quad \text{où} \quad \omega_0^2 = \omega^2 + \lambda^2\ .$$

Donc le plan complexe $z = x + i\omega^{-1}(\dot{x} + \lambda x)$ est une reparamétrisation de l'espace de phase $(x, \dot{x})$ qui rend simple le mouvement de l'oscillateur amorti.

▷ E.D.L.S. du second ordre (réelle) : $\ddot{x} + a\dot{x} + bx = 0$ ($a = 2\lambda$ et b réels)

Rappelons qu'elle a pour solution générale (si $p_1 \neq p_2$) :

$$x(t) = A_1\,e^{p_1 t} + A_2\,e^{p_2 t} \qquad ; \qquad p_1, p_2 \text{ solutions de } p^2 + ap + b = 0\ .$$

Compte tenu de la linéarité vis-à-vis des C.I. le lecteur peut établir "de tête" l'expression :

$$x(t) = x(0)\,\frac{p_2 e^{p_1 t} - p_1 e^{p_2 t}}{p_2 - p_1} + \dot{x}(0)\,\frac{e^{p_2 t} - e^{p_1 t}}{p_2 - p_1}\ .$$

Dans le cas où p_1 et p_2 sont complexes conjugués ($b > \lambda^2 > 0$ et $p_1 = -\lambda - i\omega = \overline{p_2}$), on utilise les combinaisons linéaires de $e^{-\lambda t}\cos\omega t$ et $e^{-\lambda t}\sin\omega t$:

$$x(t) = x(0)\, e^{-\lambda t}\left(\cos\omega t + \lambda\,\frac{\sin\omega t}{\omega}\right) + \dot{x}(0)\, e^{-\lambda t}\,\frac{\sin\omega t}{\omega} \qquad (\omega = \sqrt{b-\lambda^2})\ .$$

Enfin dans le cas $p_1 = p_2$ (limite $\omega \to 0$ du cas précédent) :

$$x(t) = x(0)\, e^{-\lambda t}(1+\lambda t) + \dot{x}(0)\, t\, e^{-\lambda t}\ .$$

Connaissant $x(t)$ on déduit $\dot{x}(t)$ et l'état $\underline{x}(t)$. Il s'écrit, pour $p_1 \neq p_2$:

$$\underline{x}(t) = \begin{pmatrix} x \\ \dot{x} \end{pmatrix} = A_1 e^{p_1 t}\begin{pmatrix} 1 \\ p_1 \end{pmatrix} + A_2 e^{p_2 t}\begin{pmatrix} 1 \\ p_2 \end{pmatrix} = A_1(t)\,\underline{m}_1 + A_2(t)\,\underline{m}_2\ .$$

Si p_1 et p_2 sont réels, le vecteur $\underline{x}(t)$ s'obtient par addition de deux vecteurs parallèles à $\underline{m}_1$ et $\underline{m}_2$ et d'amplitudes connues. Pour $\lambda > 0$, ces amplitudes sont décroissantes et le vecteur $\underline{x}(t)$ tend à devenir colinéaire avec le vecteur affecté de l'amplitude qui décroit la moins vite, soit $\underline{m}_1$ si $p_2 < p_1 < 0$. La trajectoire dans l'espace de phase finit donc tangentiellement à la droite $\dot{x} = p_1 x$. Une différence avec un S.D. du premier ordre est qu'il existe des C.I. telles que $x(t)$, ou $\dot{x}(t)$, change de signe en tendant vers zéro.

Si p_1 et p_2 sont complexes conjugués, il est préférable d'écrire l'état $\underline{x}(t)$ (réel), non comme la somme de deux vecteurs complexes conjugués, mais comme la partie réelle de l'un d'eux :

$$\underline{x}(t) = Ae^{-\lambda t}\,\Re e\left\{e^{-i(\omega t+\varphi)}\begin{pmatrix} 1 \\ -\lambda - i\omega \end{pmatrix}\right\} = \Re e\Big\{z(t)\,\underline{m}\Big\}\ .$$

Pour $\lambda > 0$ la trajectoire tend vers zéro en s'enroulant autour de l'origine. Les deux portraits de phase correspondant aux cas p_1 et p_2 réels ou complexes conjugués sont représentés sur les figures 9 et 10. La construction des trajectoires utilise les isoclines qui sont les droites $(\mathrm{tg}\theta + a)\dot{x} + bx = 0$.

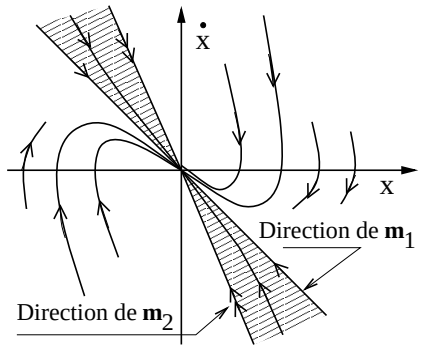


FIGURE 9

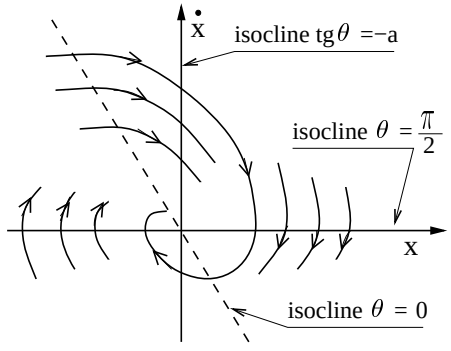


FIGURE 10

▷ E.D.L.S. avec second membre

Bien que non linéaires les équations

$$\dot{x} + \lambda\, x = f(t) \quad ; \quad \dot{z} + (\lambda + i\omega) z = f(t) \quad ; \quad \ddot{x} + a\dot{x} + bx = f(t)\ ,$$

dites "linéaires avec second membre", et qui apparaissent en physique dès qu'il y a des sources (forces en mécanique, générateurs en électricité, *etc.*), sont étudiées ici car leur résolution se ramène à l'étude précédente. La **solution générale** de ces équations dépend linéairement des C.I. et du second membre f, ce qui pour f en particulier signifie une dépendance linéaire vis-à-vis de toutes les valeurs $f(s)$ dans l'intervalle $[0, t]$. Pour les **équations du premier ordre** on vérifie :

$$\dot{x} + \lambda x = f(t) \Longrightarrow x(t) = x(0)\, e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} f(s)\, \mathrm{d}s\ ;$$

($z(t)$ est obtenu en faisant simplement les changements $x \to z$ et $\lambda \to \lambda + i\omega$). En effet la dérivée de l'intégrale $I(t)$ ci-dessus donne deux contributions : l'une obtenue en dérivant l'exponentielle vaut $-\lambda I(t)$, et l'autre en dérivant par rapport à la borne supérieure d'intégration vaut $f(t)$. Pour l'**équation du second ordre**, on vérifie de la même manière que :

$$\ddot{x} + a\dot{x} + bx = f(t) \Longrightarrow x(t) = R_0(t)\, x(0) + R_1(t)\, \dot{x}(0) + \int_0^t R_1(t-s)\, f(s)\, \mathrm{d}s\ .$$

$R_0(t)$ et $R_1(t)$ sont les solutions de l'équation sans second membre telles que $R_0(0) = 1$, $\dot{R}_0(0) = 0$ et $R_1(0) = 0,\ \dot{R}_1(0) = 1$. Les deux premiers termes constituent la solution de l'équation sans second membre satisfaisant les C.I. ; l'intégrale est la solution particulière $x_0(t)$ de l'équation avec second membre qui satisfait $x_0(0) = \dot{x}_0(0) = 0$. Exemple :

$$\ddot{x} + \omega^2 x = f(t) \Longrightarrow x(t) = x(0)\, \cos\omega t + \dot{x}(0)\, \frac{\sin\omega t}{\omega} + \int_0^t \frac{\sin\omega(t-s)}{\omega}\, f(s)\, \mathrm{d}s\ .$$

Pour l'**E.D.L. du n$^{\text{ème}}$ ordre** $x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1\dot{x} + a_0 x = f$, on peut vérifier à titre d'exercice que sa solution s'écrit (même si les coefficients dépendent du temps et qu'on ne peut pas la déterminer explicitement) :

$$x(t) = R_0(t, t_0)\, x(t_0) + R_1(t, t_0)\, \dot{x}(t_0) + \cdots + R_{n-1}(t, t_0)\, x^{(n-1)}(t_0) + \int_{t_0}^t R_{n-1}(t, s)\, f(s)\, \mathrm{d}s\ ,$$

où $R_j(t, t_0)$ est la solution de l'équation sans second membre dont toutes les C.I. (prises à t_0) sont nulles sauf la dérivée d'ordre j qui vaut 1 : $R_j^{(k)}(t_0, t_0) = \delta_{jk}\ \ (j, k = 0 \cdots n{-}1)$. Donc l'effet du second membre $f(t)$ dans tout intervalle $[s, s + \mathrm{d}s]$ est de générer une solution $R_{n-1}(t, s)\, f(s)\, \mathrm{d}s$ de l'équation sans second membre. $R_{n-1}(t, s)$ est la réponse à une impulsion $\delta(t - s)$ (on rappelle que $f(t) = \int f(s)\, \delta(t - s)\, \mathrm{d}s$).

REMARQUE. On peut aussi, pour résoudre ces équations, commencer par la recherche d'une solution particulière. Une fois une telle solution $x_0(t)$ (resp. $z_0(t)$) trouvée, il suffit de lui ajouter la solution générale de l'équation sans second membre, puisqu'en effet $x - x_0$ (resp. $z - z_0$) satisfait cette équation.

6.2.2 Cas général ; modes propres ; oscillateurs couplés

▷ Modes propres

Si on a bien pris garde de ne pas éliminer de grandeurs, l'équation générale pour un système linéaire stationnaire s'écrit

$$\dot{\underline{x}} = \mathbf{A}\,\underline{x}\ ,$$

où $\mathbf{A}$ est une matrice $n\times n$ à coefficients constants. Exemple : $\dfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ \dot{x}\end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a\end{pmatrix}\begin{pmatrix} x \\ \dot{x}\end{pmatrix}$.

La solution est facile à obtenir lorsque la matrice $\mathbf{A}$ est diagonalisable, c'est-à-dire lorsqu'il existe une base de **vecteurs propres** $\underline{m}_1, \underline{m}_2 \cdots \underline{m}_n$ satisfaisant $\mathbf{A}\,\underline{m}_\alpha = p_\alpha\,\underline{m}_\alpha$ (p_α **valeurs propres**). Si on décompose $\underline{x}(t)$ sur cette base, $\underline{x}(t) = \sum_\alpha a_\alpha(t)\,\underline{m}_\alpha$, le système de départ s'écrit $\sum_\alpha \dot{a}_\alpha\,\underline{m}_\alpha = \sum_\alpha a_\alpha\,p_\alpha\,\underline{m}_\alpha$, et se ramène donc à des E.D. du premier ordre pour chaque composante : $\dot{a}_\alpha(t) = p_\alpha\,a_\alpha(t)$. La solution générale est (A_α constantes arbitraires) :

$$\underline{x}(t) = \sum_\alpha A_\alpha\,e^{p_\alpha t}\,\underline{m}_\alpha \qquad (\mathbf{A}\,\underline{m}_\alpha = p_\alpha\,\underline{m}_\alpha)\ .$$

Chaque solution particulière $e^{p_\alpha t}\,\underline{m}_\alpha$ s'appelle un mode propre. La caractéristique d'un mode propre est que toutes les composantes $x_1(t), x_2(t)\cdots x_n(t)$ de $\underline{x}(t)$ ont le même comportement en temps, et donc ne diffèrent que par une constante multiplicative. Lorsque $\mathbf{A}$ est réelle, on distingue parmi ces modes les **modes réels non oscillants** (p et $\underline{m}$ réels) et les **modes oscillants** ($p = -(\lambda + i\omega)$ et $\underline{m} = \underline{c} + i\underline{s}$, $\underline{c}$ et $\underline{s}$ vecteurs réels). Dans le premier cas, les trajectoires dans l'espace de phase sont des demi droites de direction $\underline{m}$ passant par O. Dans le deuxième cas, comme on l'a fait pour l'équation du second ordre, on regroupe les modes complexes $\underline{m}$ et $\overline{\underline{m}}$ associés aux valeurs propres p et $\overline{p}$ complexes conjuguées sous la forme réelle :

$$A\,e^{-\lambda t}\,\Re e\left\{\underline{m}\,e^{-i(\omega t+\varphi)}\right\} = A\,e^{-\lambda t}\left(\cos(\omega t+\varphi)\,\underline{c} + \sin(\omega t+\varphi)\,\underline{s}\right)\ .$$

Les trajectoires sont alors des spirales "elliptiques" dans le plan $(\underline{c}, \underline{s})$.

REMARQUE. En présence de "forces" ou de "générateurs" extérieurs, l'équation devient :

$$\dot{\underline{x}}(t) = \mathbf{A}\,\underline{x}(t) + \underline{F}(t)\ .$$

La décomposition $\underline{F}(t) = \sum_\alpha f_\alpha(t)\,\underline{m}_\alpha$ la ramène aux n équations découplées (déjà résolues) $\dot{a}_\alpha(t) = p_\alpha\,a_\alpha(t) + f_\alpha(t) \qquad (\alpha = 1\cdots n)$. La solution peut aussi se mettre sous la forme (valable même si $\mathbf{A}$ n'est pas diagonalisable) :

$$\underline{x}(t) = e^{\mathbf{A}t}\,\underline{x}(0) + \int_0^t e^{\mathbf{A}(t-s)}\,\underline{F}(s)\,\mathrm{d}s\ .$$

EXEMPLE. **Résonance magnétique nucléaire (R.M.N.).** L'équation générale pour l'aimantation $\overrightarrow{M}(t)$ d'un milieu en présence d'un champ magnétique extérieur constant $\overrightarrow{B}_0$ et d'un champ variable $\overrightarrow{B}(t)$ (en pratique un champ tournant dans un plan perpendiculaire à $\overrightarrow{B}_0$) est :

$$\frac{\mathrm{d}\overrightarrow{M}}{\mathrm{d}t} = -\frac{\overrightarrow{M}_\parallel - \overrightarrow{M}_0}{\tau_\parallel} - \frac{\overrightarrow{M}_\perp}{\tau_\perp} + \gamma\,\overrightarrow{M}\wedge\left(\overrightarrow{B}_0 + \overrightarrow{B}(t)\right)\ ,$$

où $\overrightarrow{M}_{\parallel}$ et $\overrightarrow{M}_{\perp}$ désignent respectivement les composantes de $\overrightarrow{M}(t)$ parallèle et perpendiculaire à $\overrightarrow{B}_0$, et où $\overrightarrow{M}_0$ est un vecteur constant parallèle à $\overrightarrow{B}_0$. Pour $\overrightarrow{B}(t) = 0$, cette équation vectorielle s'écrit

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} M_1 \\ M_2 \\ M_3 - M_0 \end{pmatrix} = \begin{pmatrix} -\lambda_\perp & -\omega_0 & 0 \\ \omega_0 & -\lambda_\perp & 0 \\ 0 & 0 & -\lambda_\parallel \end{pmatrix}\begin{pmatrix} M_1 \\ M_2 \\ M_3 - M_0 \end{pmatrix} \quad \text{avec } \lambda_{\perp,\parallel} = \frac{1}{\tau_{\perp,\parallel}} \text{ et } \omega_0 = -\gamma B_0 \ ,$$

l'axe 3 étant choisi selon $\overrightarrow{B}_0$ ($M_3 \equiv M_\parallel$). C'est l'équation d'un S.D.L. à trois degrés de liberté, qui possède un mode propre réel correspondant à la relaxation sans oscillation (temps caractéristique $\tau_\parallel$) de la composante $M_\parallel$ vers sa valeur d'équilibre M_0. Les deux autres modes sont complexes conjugués et correspondent à la relaxation de $\overrightarrow{M}_\perp$ vers zéro (temps caractéristique $\tau_\perp$), accompagnée d'une rotation à la vitesse angulaire ω_0 (figure 11) ; en effet $z = M_1 + iM_2$ vérifie $\dot{z} = (-\lambda_\perp + i\omega_0)z$.

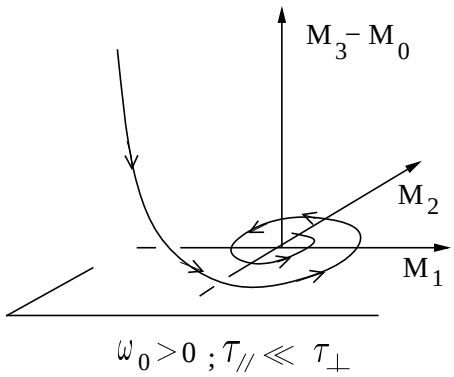


FIGURE 11

La présence de $\overrightarrow{B}(t)$ de composantes $B_1(t) = B\cos\omega t$ et $B_2(t) = B\sin\omega t$ a pour effet principal d'induire le phénomène de résonance pour $\overrightarrow{M}_\perp$. Considérons pour simplifier le cas où le terme en $\overrightarrow{M}_\perp \wedge \overrightarrow{B}(t)$ est négligeable. L'évolution de $\overrightarrow{M}_\parallel(t)$ n'est alors pas modifiée ; par contre l'équation pour $\overrightarrow{M}_\perp$, c.a.d. pour z, devient :

$$\dot{z} = (-\lambda_\perp + i\omega_0)\, z + i\gamma M_3 B e^{i\omega t} \ .$$

En régime permanent, donc pour $t \gg \tau_\parallel$, on peut remplacer $M_3(t)$ par sa valeur limite M_0. On obtient la solution stationnaire :

$$z(t) = |z|\, e^{i\omega t} \quad ; \quad |z| = \gamma M_0 B \left(\lambda_\perp^2 + (\omega - \omega_0)^2\right)^{-\frac{1}{2}} = |\overrightarrow{M}_\perp| \ .$$

$|\overrightarrow{M}_\perp|$ présente une résonance pour $\omega = \omega_0$ (quantité algébrique), c'est-à-dire lorsque $\overrightarrow{B}(t)$ tourne dans le même sens et à la même vitesse que $\overrightarrow{M}$ en absence de champ tournant.

▷ Cas des oscillateurs couplés sans frottement

Si $\frac{1}{2}\,\underline{\dot{x}}^t \mathbf{M} \underline{\dot{x}}$ et $\frac{1}{2}\,\underline{x}^t \mathbf{K} \underline{x}$ sont les formes quadratiques définies positives exprimant l'énergie cinétique et l'énergie potentielle en fonction des écarts $x_1(t), x_2(t) \cdots$ à l'équilibre, l'équation du mouvement en présence de forces extérieures a la forme générale (*cf.* section 4.2.2) :

$$\mathbf{M}\,\underline{\ddot{x}} = -\mathbf{K}\,\underline{x} + \underline{F} \quad \text{avec } \mathbf{M} \text{ et } \mathbf{K} \text{ matrices symétriques positives .}$$

Par exemple pour le système de masses et de ressorts de la figure 12, les équations sont

$m\ddot{x}_1 = -2kx_1 + kx_2 + F_1$ et $m\ddot{x}_2 = -2kx_2 + kx_1 + F_2$ ou :

$$m\begin{pmatrix}\ddot{x}_1\\ \ddot{x}_2\end{pmatrix} = -\begin{pmatrix}2k & -k\\ -k & 2k\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + \begin{pmatrix}F_1\\ F_2\end{pmatrix} \ ;$$

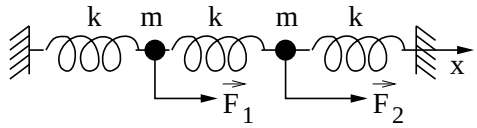


FIGURE 12

le système non couplé est obtenu en ne gardant que la partie diagonale de $\mathbf{K}$. Plutôt que de passer à des équations du premier ordre comme on l'a fait jusqu'à présent, il est plus simple d'introduire les solutions de $\mathbf{M}\,\underline{\ddot{x}} = -\mathbf{K}\,\underline{x}$ de la forme $\underline{x}(t) = e^{p_\alpha t}\underline{m}_\alpha$. Elles doivent vérifier

$$p_\alpha^2\,\mathbf{M}\,\underline{m}_\alpha = -\mathbf{K}\,\underline{m}_\alpha \ ,$$

ce qui implique que les p_α sont solutions de $\det\big(\mathbf{K}+p_\alpha^2\mathbf{M}\big) = 0$. On a vu que les propriétés de $\mathbf{M}$ et $\mathbf{K}$ entraînent que les $\underline{m}_\alpha$ forment une base réelle, et que les p_α^2 sont du type $p_\alpha^2 = -\omega_\alpha^2$. En écrivant $\underline{x}(t) = \sum_\alpha a_\alpha(t)\,\underline{m}_\alpha$ et $\underline{F} = \sum_\alpha f_\alpha(t)\,\mathbf{M}\,\underline{m}_\alpha$, on obtient comme précédemment n E.D. découplées, mais du second ordre, pour les nouvelles composantes $a_\alpha(t)$ de $\underline{x}(t)$:

$$\ddot{a}_\alpha(t) = -\omega_\alpha^2\,a_\alpha(t) + \ f_\alpha(t) \ .$$

Donc un ensemble d'oscillateurs couplés $x_1(t), x_2(t)\cdots$ se ramène à un ensemble d'oscillateurs indépendants $a_1(t), a_2(t)\cdots$ à condition de choisir les bons vecteurs de base, ou ce qui est équivalent, les bonnes combinaisons linéaires des $x_i(t)$. La **solution générale** pour $\underline{F} = 0$

$$\underline{x}(t) = \sum_\alpha A_\alpha\, cos(\omega_\alpha t + \varphi_\alpha)\,\underline{m}_\alpha$$

se présente comme une somme de n solutions appelées **modes propres d'oscillation** de pulsations ω_α (à distinguer des modes propres introduits précédemment qui sont en nombre double). Dans un tel mode, associé au vecteur $\underline{m}_\alpha$, toutes les variables $x_1(t), x_2(t)\cdots$ oscillent en phase, ou en opposition de phase (suivant les signes relatifs des différentes composantes du vecteur $\underline{m}_\alpha$), avec la même pulsation ω_α.

En présence de forces extérieures sinusoïdales $\underline{F}(t) = \underline{F}\cos\omega t$, on observe une **résonance** lorsque ω est égal à l'une des pulsations propres ω_α. Si par exemple $\omega = \omega_1$, toutes les amplitudes $a_\alpha(t)$ $(\alpha \neq 1)$ restent finies alors que $a_1(t)$ tend vers l'infini, et on a par conséquent $\underline{x}(t) \simeq a_1(t)\,\underline{m}_1$: une excitation à la pulsation propre ω_α met en évidence le mode propre "α" correspondant.

Dans l'exemple ci-dessus on a :

$$\underline{m}_1 = \begin{pmatrix}1\\1\end{pmatrix} \ , \ \omega_1^2 = \frac{k}{m} \ \ ; \ \ \underline{m}_2 = \begin{pmatrix}1\\-1\end{pmatrix} \ , \ \omega_2^2 = \frac{3k}{m} \ \ ; \ \ \underline{F} = \frac{F_1+F_2}{2}\,\underline{m}_1 + \frac{F_1-F_2}{2}\,\underline{m}_2 \ .$$

Le mode 1 (resp. 2) rentre en résonance si $\omega = \omega_1$ (resp. ω_2), sauf si $F_1 = -F_2$ (resp. $F_1 = F_2$). Les bonnes combinaisons linéaires de x_1 et x_2 sont dans ce cas $x_1 \pm x_2$:

$$m(\ddot{x}_1 + \ddot{x}_2) = -k(x_1 + x_2) + F_1 + F_2 \quad \text{et} \quad m(\ddot{x}_1 - \ddot{x}_2) = -3k(x_1 - x_2) + F_1 - F_2 \ .$$

$\dfrac{x_1 + x_2}{2} = a_1$ et $\dfrac{x_1 - x_2}{2} = a_2$ sont les composantes de $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ dans la "bonne" base $\underline{m}_1, \underline{m}_2$.

6.2.3 Stabilité et instabilité d'un système dynamique linéaire stationnaire

▷ Conditions de stabilité

Un S.D.L.S. est stable si toute solution tend vers la valeur d'équilibre $\underline{x} = 0$. Toute solution étant une combinaison de modes propres, ceci implique que tous les modes (solutions dont le comportement est en $e^{p_\alpha t}$) soient stables, donc que :

$$\Re e\, p_\alpha < 0 \quad \forall \alpha \ .$$

EXEMPLE : **stabilité d'un S.D.L.S. à deux degrés de liberté**. Suivant l'écriture des équations, les conditions de stabilité sont :

E.D. du second ordre	forme matricielle	forme complexe
$\ddot{x} + a\dot{x} + bx = 0$	$\dfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{M}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$	$\dot{z} = \alpha z + \beta \overline{z}$
$a > 0\ ,\ b > 0$	$\mathrm{tr}\,\mathbf{M} < 0\ ,\ \det \mathbf{M} > 0$	$\Re e\, \alpha < 0\ ,\ \lvert\alpha\rvert > \lvert\beta\rvert$

Démonstration : dans le cas de l'E.D. du second ordre, on a $p^2 + ap + b = 0$ donc $p_1 + p_2 = -a$ et $p_1 p_2 = b$; la condition $a > 0$ et $b > 0$ équivaut à, soit p_1 et p_2 réels négatifs, soit $p_{1,2} = -\lambda \pm i\omega$ complexes conjugués avec une partie réelle négative. Dans le cas matriciel p_1 et p_2 sont les valeurs propres de $\mathbf{M}$, donc $p_1 + p_2 = \mathrm{tr}\,\mathbf{M}$ et $p_1 p_2 = \det \mathbf{M}$. Enfin l'E.D. complexe avec $z = x_1 + ix_2$, $\alpha = \alpha_1 + i\alpha_2$ et $\beta = \beta_1 + i\beta_2$ est équivalente à une E.D. matricielle pour x_1, x_2 avec $\mathbf{M} = \begin{pmatrix} \alpha_1 + \beta_1 & \beta_2 - \alpha_2 \\ \beta_2 + \alpha_2 & \alpha_1 - \beta_1 \end{pmatrix}$.

▷ Bifurcations vers l'instabilité (figure 13)

Quand, faisant varier un paramètre, un système stable devient instable, en général quatre cas peuvent se produire :
- soit une valeur propre p réelle négative devient positive (exemple 1) ;
- soit deux valeurs propres complexes conjuguées $-\lambda \pm i\omega$ ont leur partie réelle négative qui devient positive (exemple 2) ;
- soit un mode oscillant disparaît (exemple 3) ;
- soit les fréquences initialement distinctes de deux modes oscillants deviennent égales (exemple 4).

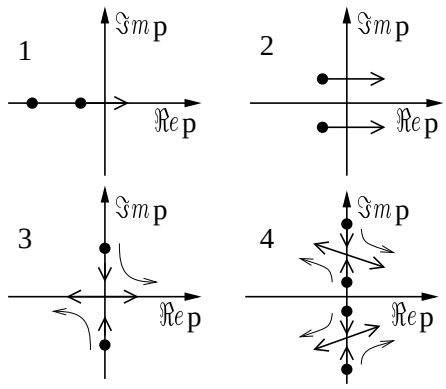


FIGURE 13

EXEMPLE 1. L'équation $\ddot{x} + a\dot{x} + bx = 0$ avec $a > 0$ voit un de ses modes non oscillants stables devenir instable (non oscillant) lorsque b passe d'une valeur positive à une valeur négative.

EXEMPLE 2. La même équation avec $b > 0$ voit son mode oscillant stable devenir instable (oscillant) lorsque a ("coefficient de frottement") passe d'une valeur positive à une valeur négative.

EXEMPLE 3. L'équation $\ddot{x} + bx = 0$ voit son mode oscillant marginalement stable donner naissance à deux modes non oscillants, dont un instable, lorsque b passe d'une valeur positive à une valeur négative (passage d'un minimum à un maximum d'énergie potentielle).

EXEMPLE 4. **Système à deux modes oscillants** : $\dddot{x} + a\ddot{x} + bx = 0$. Un exemple est fourni par le système des oscillateurs couplés de la figure 14

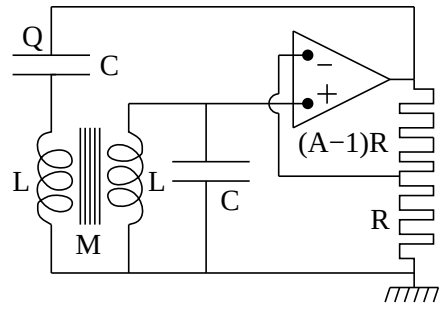


FIGURE 14

qui est régi par l'équation

$$(L^2 - M^2)\,\dddot{Q} + \frac{2L - AM}{C}\,\ddot{Q} + \frac{Q}{C^2} = 0\ ,$$

ou sous forme adimensionnée :

$$(1-m^2)\dddot{x} + (2-mA)\ddot{x} + x = 0 \ ;$$

$m = \dfrac{M}{L}$ est tel que $|m| < 1$. Ce système est marginalement stable pour $a > 0$, $b > 0$ et $a^2 > 4b$; en effet il possède deux modes oscillants car l'équation $p^4 + ap^2 + b = 0$ a quatre racines imaginaires pures conjuguées deux à deux $p_1 = -\overline{p}_1 = i\omega_1$ et $p_2 = -\overline{p}_2 = i\omega_2$. Si, lorsque les paramètres a et b (ou A et m) varient, ω_1 tend vers ω_2, c'est que dans le plan (a, b) le point représentatif P se rapproche d'un point Q de la courbe d'équation $a^2 = 4b$ (figure 15). Si cette courbe est franchie (point R dans le plan (a, b)), p_1^2 et p_2^2 deviennent complexes, et les quatre racines $\pm\sqrt{p_1^2}$ et $\pm\sqrt{p_2^2}$ ont deux à deux des parties réelles opposées. Un des modes oscillants devient stable et l'autre instable.

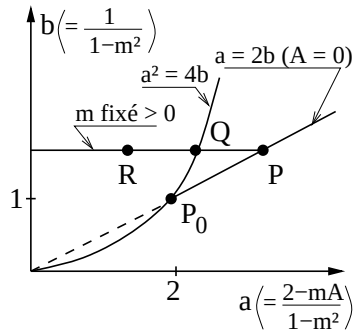


FIGURE 15

Remarque. Cette **instabilité par confusion de fréquences** ne peut se produire dans la réalité que si le système reçoit de l'énergie de l'extérieur. La condition d'instabilité $(2-mA)^2 < 4(1-m^2)$ n'est jamais satisfaite sans amplificateur opérationnel. Si $A = 0$ le système des oscillateurs couplés par la mutuelle M reste marginalement stable : point P dans le plan (a, b). Certes les fréquences propres se confondent pour $m = 0$, mais la courbe $a^2 = 4b$ n'est jamais franchie. L'instabilité par confusion de fréquences des modes de torsion et de flexion d'ailes d'avion est possible car le couplage de ces deux modes se fait avec apport d'énergie lié au mouvement de translation dans l'air avec frottement ; de même l'instabilité à l'origine du rayonnement laser correspond à la confusion d'une fréquence de la cavité et d'une fréquence atomique avec apport d'énergie lié au pompage optique.

6.3 DIX ÉQUATIONS VECTORIELLES CLASSIQUES

Un certain nombre de mouvements classiques faisant partie de la culture du physicien sont brièvement étudiés ici. Certains ont déjà été vu dans les sections 2.3.3 et 2.3.4.

▷ Mouvement uniformément accéléré : $\ddot{\vec{r}} = \vec{a}$

La solution

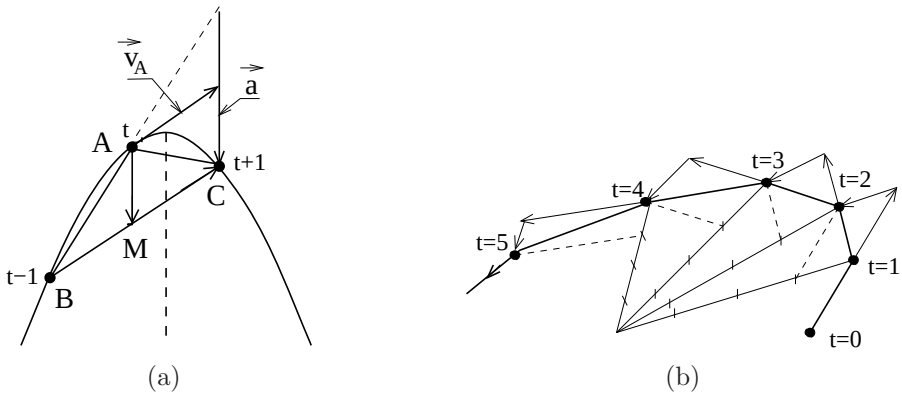
$$\vec{r}(t) = \vec{r}_0 + \vec{v}_0 t + \frac{1}{2}\vec{a}t^2 \quad \text{où} \quad \vec{r}(0) = \vec{r}_0 \ , \ \dot{\vec{r}}(0) = \vec{v}_0$$

correspond à une trajectoire parabolique d'axe parallèle à $\vec{a}$ et située dans le plan passant par $\vec{r}_0$ et parallèle à $\vec{v}_0$ et $\vec{a}$. Une propriété caractéristique moins connue de ce mouvement

est la possibilité d'exprimer exactement la vitesse et l'accélération à l'aide d'accroissements finis

$$\vec{v}(t) = \frac{\vec{r}(t+T) - \vec{r}(t-T)}{2T} \quad , \quad \vec{a}(t) = \frac{\vec{r}(t+T) + \vec{r}(t-T) - 2\vec{r}(t)}{T^2} \ ,$$

et donc d'avoir une construction géométrique précise de ces grandeurs : sur la figure 16a, $\vec{v}_A = \dfrac{\overrightarrow{BC}}{2}$ et $\vec{a} = 2\overrightarrow{AM}$ $(T = 1)$. Pour des mouvements quelconques l'expression de $\vec{a}$ reste souvent une bonne approximation ; elle permet de déduire la position $\vec{r}(t+T)$ à partir de $\vec{r}(t-T)$, $\vec{r}(t)$ et $\vec{a}(t)$; la figure 16b fournit par exemple la résolution graphique approchée de $\ddot{\vec{r}} = -\dfrac{\vec{r}}{4}$ connaissant $\vec{r}(0)$ et $\vec{r}(1)$.



(a) (b)

FIGURE 16

Une autre propriété utilisée en optique électronique, spectrographie de masse... est illustrée sur la figure 17. Si une particule pénètre à $t = 0$ en I dans un champ d'accélération uniforme avec une vitesse $\vec{v}_I \perp \vec{a}$, et le quitte en J à l'instant t, la relation

$$\overrightarrow{IJ} = \vec{v}_I t + \tfrac{1}{2}\vec{a}t^2 = \tfrac{1}{2}\left(\vec{v}_I t + \vec{v}_J t\right) = \tfrac{\overrightarrow{IK}}{2} + \tfrac{1}{2}\vec{v}_J t$$

montre que la trajectoire, rectiligne à la sortie du champ, semble venir du milieu M du segment IK, quel que soit le module $|\vec{v}_I|$ de la vitesse initiale.

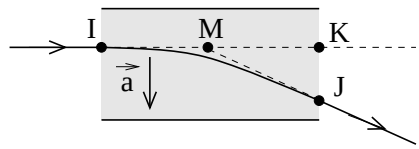


FIGURE 17

REMARQUES 1) En présence de **frottement visqueux**, l'équation devient $\dot{\vec{v}} + \frac{\vec{v}}{\tau} = \vec{a}$ et s'intègre facilement : $\vec{v}(t) = \vec{a}\tau + (\vec{v}_0 - \vec{a}\tau)\, e^{-\frac{t}{\tau}}$ et $\vec{r}(t) = \vec{r}_0 + \vec{a}\tau t - (\vec{v}_0 - \vec{a}\tau)\,\tau\,(e^{-\frac{t}{\tau}} - 1)$. La trajectoire présente une asymptote parallèle à $\vec{a}$ et passant par le point de rayon vecteur $\vec{r}_0 + \vec{v}_0\tau$.

2) L'équation de la **chute libre dans le référentiel local** est $\dot{\vec{v}} = \vec{g} - 2\vec{\omega} \wedge \vec{v}$ (*cf.* section 3.3.1). Elle

se résoud, au premier ordre en ω, en posant $\vec{v}(t) = \vec{v}_0 + \vec{g}t$ dans le membre de droite. On obtient :

$$\vec{r}(t) = \vec{r}_0 + \vec{v}_0 t + \tfrac{1}{2}\vec{g}t^2 - \overrightarrow{\omega} \wedge \left(\vec{v}_0 t^2 + \vec{g}\tfrac{t^3}{3}\right) .$$

Pour $\vec{v}_0 = 0$, la correction $-\overrightarrow{\omega} \wedge \vec{g}\,\frac{t^3}{3}$ décrit la "déviation vers l'est" (due à la **force de Coriolis**).

▷ Mouvement harmonique : $\ddot{\vec{r}} + \omega^2 \vec{r} = 0 \quad \left(\omega^2 = \dfrac{k}{m}\right)$

Sa solution

$$\vec{r}(t) = \vec{r}_0 \cos\omega t + \vec{v}_0 \frac{\sin\omega t}{\omega} \quad \text{où} \quad \vec{r}(0) = \vec{r}_0 \ , \ \dot{\vec{r}}(0) = \vec{v}_0$$

a été étudiée à la section 2.3.4. On vérifie aisément qu'elle possède neuf intégrales premières (grandeurs conservées) $\dot{x}^2 + \omega^2 x^2$, $\dot{y}^2 + \omega^2 y^2$, $\dot{z}^2 + \omega^2 z^2$, $\dot{x}\dot{y} + \omega^2 xy$, $\dot{x}\dot{z} + \omega^2 xz$, $\dot{y}\dot{z} + \omega^2 yz$, $x\dot{y} - y\dot{x}$, $x\dot{z} - z\dot{x}$, et $y\dot{z} - z\dot{y}$, soit cinq de plus que les quatre attendues que sont l'énergie $E = \dfrac{1}{2}(m\dot{\vec{r}}^2 + k\vec{r}^2)$ et les trois composantes du moment cinétique $\overrightarrow{L} = \vec{r} \wedge m\dot{\vec{r}}$.

▷ Mouvement de Kepler : $\ddot{\vec{r}} = -\alpha \dfrac{\vec{r}}{r^3}$

Ce mouvement étudié à la section 2.3.3. possède les intégrales premières :

$$\text{“}E\text{”} = \frac{1}{2}\dot{\vec{r}}^2 - \frac{\alpha}{r} \quad ; \quad \overrightarrow{C} = \vec{r} \wedge \dot{\vec{r}} \quad ; \quad \overrightarrow{A} = \dot{\vec{r}} \wedge \overrightarrow{C} - \alpha\frac{\vec{r}}{r} \ .$$

“E” est ici proportionnel à l'énergie totale et $\overrightarrow{C}$ (constante des aires) au moment cinétique. $\overrightarrow{A}$ est le **vecteur de Lenz**. (On vérifie que $\dfrac{\mathrm{d}\overrightarrow{A}}{\mathrm{d}t} = \ddot{\vec{r}} \wedge (\vec{r} \wedge \dot{\vec{r}}) - \alpha\dfrac{\dot{\vec{r}}}{r} + \alpha\dfrac{\dot{r}\vec{r}}{r^2} = 0$ en utilisant l'équation du mouvement pour écrire le premier terme sous la forme $-\alpha\dfrac{\vec{r}}{r^3} \wedge (\vec{r} \wedge \dot{\vec{r}}) = -\alpha\dfrac{\vec{r}\cdot\dot{\vec{r}}}{r^3}\vec{r} + \alpha\dfrac{\dot{\vec{r}}}{r}$, et en utilisant l'identité $r\dot{r} = \dfrac{1}{2}\dfrac{\mathrm{d}\vec{r}^2}{\mathrm{d}t} = \vec{r}\cdot\dot{\vec{r}}$ pour écrire le troisième terme sous la forme $\alpha\dfrac{\vec{r}(\vec{r}\cdot\dot{\vec{r}})}{r^3}$). $\vec{r}$ étant orthogonal à $\overrightarrow{C}$, la trajectoire est dans le plan passant par O et perpendiculaire à $\overrightarrow{C}$; $\overrightarrow{A}$ est situé dans ce plan. Les relations

$$\overrightarrow{A}\cdot\vec{r} = \overrightarrow{C}^2 - \alpha r \quad \text{et} \quad \overrightarrow{A}^2 = \dot{\vec{r}}^2\overrightarrow{C}^2 + \alpha^2 - 2\alpha\frac{\overrightarrow{C}^2}{r} = \alpha^2 + 2\text{“}E\text{”}\overrightarrow{C}^2$$

montrent que, avec un choix convenable de l'origine des angles, l'équation de la trajectoire en coordonnées polaires s'écrit :

$$r = \frac{p}{1 + e\cos\theta} \qquad \text{avec} \qquad p = \frac{\overrightarrow{C}^2}{\alpha} \ , \ e = \frac{|\overrightarrow{A}|}{|\alpha|} = \sqrt{1 + \frac{2\text{“}E\text{”}\overrightarrow{C}^2}{\alpha^2}} \ .$$

Il s'agit donc, comme on l'a déjà vu, d'ellipses si $e < 1$ donc pour “E” < 0 (ce qui implique $\alpha > 0$), de paraboles si $e = 1$ donc pour “$E = 0$” ($\alpha > 0$ aussi), et de branches

d'hyperboles si $e > 1$ donc pour "E" > 0 (α quelconque). On retrouve l'angle Δ de la diffusion de Rutherford (*cf.* section 2.3.3) en écrivant la conservation de $\overrightarrow{A}$ à $t = \pm\infty$:

$$\overrightarrow{A} = \vec{v}_1 \wedge \overrightarrow{C} - \alpha\hat{r}_1 = \vec{v}_2 \wedge \overrightarrow{C} - \alpha\hat{r}_2 \Longrightarrow (\vec{v}_2 - \vec{v}_1) \wedge \overrightarrow{C} = \alpha(\hat{r}_2 - \hat{r}_1) \ .$$

▷ Mouvements impliquant des vecteurs tournants

Rappelons l'équation satisfaite par un vecteur $\overrightarrow{V}(t)$ qui tourne à vitesse angulaire $\overrightarrow{\omega}$ uniforme, et l'expression de $\overrightarrow{V}(t)$ en fonction des valeurs initiales de ses composantes parallèle $\overrightarrow{V}_{0\|}$ et perpendiculaire $\overrightarrow{V}_{0\perp}$ à $\overrightarrow{\omega}$:

$$\frac{\mathrm{d}\overrightarrow{V}(t)}{\mathrm{d}t} = \overrightarrow{\omega} \wedge \overrightarrow{V}(t) \quad ; \quad \overrightarrow{V}(t) = \overrightarrow{V}_{0\|} + \cos\omega t\, \overrightarrow{V}_{0\perp} + \frac{\sin\omega t}{\omega}\ \overrightarrow{\omega} \wedge \overrightarrow{V}_{0\perp} \ .$$

Les exemples physiques sont multiples.

- **Mouvement de précession d'une toupie** (figure 18).

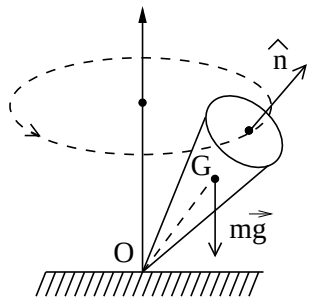


FIGURE 18

Dans l'approximation gyroscopique, le moment cinétique est $\overrightarrow{J} = I\Omega\hat{n}$, où $\overrightarrow{\Omega} = \Omega\,\hat{n}$ est le vecteur rotation de la toupie et I le moment d'inertie selon l'axe $\overrightarrow{OG}$; O est le point de contact, supposé fixe, avec le sol et G le c.d.m. On déduit du théorème du moment cinétique

$$\frac{\mathrm{d}\overrightarrow{J}}{\mathrm{d}t} = l\,\hat{n}(t) \wedge m\vec{g} \quad (l = OG) \ ,$$

d'abord que $|\overrightarrow{J}| = I\Omega$ est constant ($\overrightarrow{J} \cdot \dfrac{\mathrm{d}\overrightarrow{J}}{\mathrm{d}t} = 0$), puis que l'axe de la toupie (parallèle à $\hat{n}$) précesse à la vitesse angulaire $\overrightarrow{\omega} = -\dfrac{ml}{I\Omega}\,\vec{g}$ autour de la verticale. L'approximation est valable si $\omega \ll \Omega$.

- **Précession d'un moment magnétique $\vec{\mu}$ dans un champ magnétique $\overrightarrow{B}$.** ($\vec{\mu} = \gamma_g\,\overrightarrow{J}$ où γ_g est le facteur gyromagnétique). L'équation $\dfrac{\mathrm{d}\overrightarrow{J}}{\mathrm{d}t} = \dfrac{1}{\gamma_g}\dfrac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} = \vec{\mu} \wedge \overrightarrow{B}$ montre que $\vec{\mu}$ précesse autour de $\overrightarrow{B}$ à la vitesse angulaire $\vec{\omega} = -\gamma_g\,\overrightarrow{B}$.

- **Mouvement d'une particule chargée, relativiste ou non, dans un champ** $\overrightarrow{B}$ (figure 19). La conservation de l'énergie $E = \gamma mc^2 = \left(1 - \dfrac{v^2}{c^2}\right)^{-\frac{1}{2}} mc^2$ (ou $\dfrac{1}{2}mv^2$), conséquence de $\vec{v} \cdot \mathrm{d}\vec{p} = 0$, implique celle du module de la quantité de mouvement $\vec{p} = \gamma m\vec{v}$ (ou $m\vec{v}$). L'équation

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q\vec{v} \wedge \overrightarrow{B} = \gamma m \frac{\mathrm{d}\vec{v}}{\mathrm{d}t}$$

montre que la vitesse $\vec{v}$ précesse autour de $\overrightarrow{B}$ à la vitesse angulaire $\vec{\omega} = -\dfrac{q}{m\gamma}\overrightarrow{B}$ (ou $-\dfrac{q}{m}\overrightarrow{B}$). La trajectoire $\vec{r}(t)$ s'obtient immédiatement par intégration de $\vec{v}(t)$:

$$\vec{r}(t) = \vec{r}_0 + \vec{v}_{0\parallel}\, t + \frac{\sin \omega t}{\omega}\, \vec{v}_{0\perp} + \frac{1 - \cos \omega t}{\omega^2}\, \vec{\omega} \wedge \vec{v}_{0\perp} \ .$$

On voit sur cette expression que le mouvement transversal se fait sur un cercle centré en $\vec{r}_0 + \dfrac{\vec{\omega} \wedge \vec{v}_{0\perp}}{\omega^2}$ et de rayon $\dfrac{v_{0\perp}}{\omega}$, dans le sens direct par rapport à $\overrightarrow{B}$ si $q < 0$, et indirect si $q > 0$. La trajectoire est une hélice d'axe parallèle à $\overrightarrow{B}$ et de pas $\dfrac{2\pi}{\omega}\, \vec{v}_{0\parallel}$ (représentée sur la figure pour des charges opposées et des C.I. identiques).

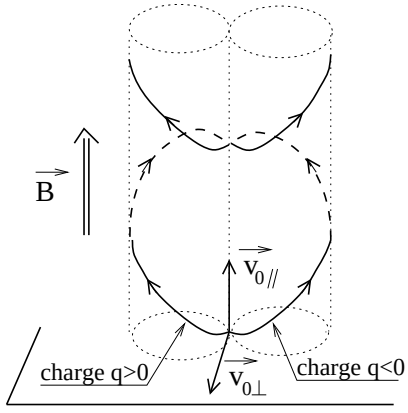


FIGURE 19

- **Pendule de Foucault et effet Zeeman**. L'équation du pendule de Foucault est :

$$m\,\ddot{\vec{r}} = -m\omega_0^2 \vec{r} - 2m\vec{\omega}_{_T} \wedge \dot{\vec{r}}\ .$$

$\omega_0 = \sqrt{g/l}$ est la pulsation du pendule ; $\overrightarrow{\omega}_{_T} = \omega \cos\theta\, \hat{z}$ est la composante verticale locale de la vitesse de rotation $\overrightarrow{\omega}$ de la Terre. Cette équation apparaît comme un cas particulier bidimensionnel, le déplacement du pendule $\vec{r}(t)$ s'effectuant dans le plan horizontal xOy, de l'équation tridimensionnelle

$$m\,\ddot{\vec{r}} = -m\omega_0^2 \vec{r} + q\dot{\vec{r}} \wedge \overrightarrow{B}\ ,$$

qui décrit le mouvement non amorti d'un électron atomique galiléen en présence d'un champ magnétique $\overrightarrow{B}$ (selon Oz). En posant $Z = x + iy$ cette dernière équation se ramène à :

$$\ddot{Z} = -\omega_0^2\, Z + i\omega_c \dot{Z} \quad \text{et} \quad \ddot{z} = -\omega_0^2\ z\ ,$$

où $\omega_c = -\frac{qB}{m}$ est la pulsation cyclotron. Le mouvement selon z est celui d'un oscillateur harmonique.

Quant au mouvement transverse il résulte, en vertu de la linéarité de l'équation, de l'addition des vecteurs tournants

$$Z_+(t) = a\,e^{i\omega_+ t} \quad \text{et} \quad Z_-(t) = b\,e^{i\omega_- t}\ ,$$

où $\omega_\pm$ sont les racines positive et négative de l'équation $\omega^2 - \omega\omega_c - \omega_0^2 = 0$. Le premier tourne dans le sens direct à la pulsation $\omega_+ \simeq \omega_0 + \frac{\omega_c}{2}$ (si $|\omega_c| \ll \omega_0$), le second dans le sens rétrograde à la pulsation $\omega_- \simeq -\omega_0 + \frac{\omega_c}{2}$. L'effet Zeeman est la traduction expérimentale de ces mouvements sur les champs rayonnés par la charge : existence de trois fréquences ω_0 et $|\omega_\pm|$, avec leurs polarisations respectivement rectiligne et elliptiques obtenues par projection du mouvement sur un plan perpendiculaire à la direction de propagation (figure 20).

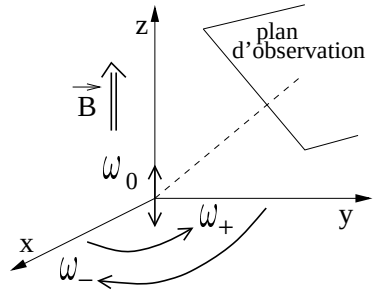


FIGURE 20

Pour le pendule de Foucault seul le mouvement transverse existe. Les solutions $Z_\pm(t)$ sont valables, mais correspondent à des conditions de lancement du pendule très particulières. La solution générale

$$Z(t) = a\,e^{i\omega_+ t} + b\,e^{i\omega_- t} \simeq (a\,e^{i\omega_0 t} + b\,e^{-i\omega_0 t})\,e^{i\frac{\omega_c}{2}t}$$

correspond à la précession lente (due à la force de Coriolis) dans le sens rétrograde, à la pulsation $\frac{\omega_c}{2} = -\omega_T = -\omega\cos\theta$, d'un mouvement elliptique (rectiligne si $|a| = |b|$; *cf.* section 2.3.4) dans le plan (xOy).

6.4 ÉQUATIONS DIFFÉRENTIELLES À COEFFICIENTS VARIABLES

Le but de ce paragraphe est d'exploiter la linéarité d'une E.D. du second ordre pour décrire aussi bien : la conjugaison objet-image en optique de Gauss, la quantification des niveaux d'énergie en mécanique quantique, ou celle des fréquences des ondes stationnaires en mécanique classique, la structure en bandes des niveaux d'énergie électroniques dans un cristal, ou celle des fréquences de résonance d'un oscillateur paramétrique. Il est aussi de montrer que, lorsque la variable est l'espace et non le temps, les **conditions aux limites** (**C.L.**) jouent souvent un rôle plus important que les C.I.

6.4.1 Quatre exemples ; matrices de transfert

▷ Exemples : mécanique ; optique ; quantique ; ondes

Considérons les équations :

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \omega^2(t)\,x = 0 \qquad \text{et} \qquad \frac{\mathrm{d}}{\mathrm{d}z}\left(\alpha(z)\,\frac{dx}{dz}\right) = -\beta(z)\,x\ .$$

La première décrit le mouvement $x(t)$ d'un **oscillateur paramétrique**, dont la pulsation

$\omega(t)$ varie. La seconde est une récriture de l'équation du second ordre la plus générale. Elle correspond par exemple à la **trajectoire** $x(z)$ dans le plan xOz **d'un rayon lumineux** "proche" de l'axe Oz d'une fibre optique d'indice $n(x,y,z) \simeq \alpha(z) - \beta(z)\,\dfrac{x^2+y^2}{2}$ (approximation linéaire de la loi de Descartes $\dfrac{\mathrm{d}}{\mathrm{d}l}(n\hat{u}) = \overrightarrow{\mathrm{grad}}\, n$ obtenue en posant $\mathrm{d}l \simeq \mathrm{d}z$, $u_x \simeq \dfrac{\mathrm{d}x}{\mathrm{d}z}$), ou encore à des trajectoires électroniques près de l'axe dans un potentiel axisymétrique.

Considérons de même les équations d'ondes à une dimension :

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}z^2} + \big(V(z) - E\big)\,\psi = 0 \qquad \text{et} \qquad \frac{\mathrm{d}}{\mathrm{d}z}\left(\alpha(z)\,\frac{\mathrm{d}\psi}{\mathrm{d}z}\right) = -\beta(z,\omega)\,\psi\ .$$

La première décrit un **état stationnaire quantique** $\Psi(t,z) = e^{-i\frac{Et}{\hbar}}\psi(z)$ d'une particule d'énergie potentielle $V(z)$ (E.D. de Schrödinger), et la seconde décrit des **ondes classiques** $\Psi(t,z) = e^{-i\omega t}\psi(z)$ en notation complexe. Par exemple, l'équation des ondes sonores dans un fluide inhomogène de masse volumique $\rho(z)$ et de compressibilité $\chi(z)$

$$\rho(z)\frac{\partial^2\Psi}{\partial t^2} = -\frac{\partial P'}{\partial z} \quad \text{avec } P' \text{ (surpression)} = -\frac{1}{\chi(z)}\,\frac{\partial\Psi}{\partial z}$$

conduit à une équation de cette forme pour $\psi(z)$ avec $\alpha = \chi^{-1}$ et $\beta = \rho\,\omega^2$. D'autres exemples sont fournis par les cordes vibrantes, les ondes électromagnétiques guidées, et aussi par bon nombre de problèmes classiques ou quantiques où les variables radiale et angulaires peuvent être séparées (*cf.* section 8.3.2).

▷ Matrice de transfert

Toutes ces équations recouvrent des réalités physiques très différentes, mais leur contenu mathématique est essentiellement le même. C'est pourquoi, afin d'éviter d'inutiles répétitions, nous énonçons les principaux résultats pour l'équation relative à $x(z)$, résultats que nous transposerons ensuite aux autres exemples. Cette équation s'écrit, comme pour un S.D.L.,

$$\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix}x(z)\\p(z)\end{pmatrix} = \begin{pmatrix}0 & (\alpha(z))^{-1}\\ -\beta(z) & 0\end{pmatrix}\begin{pmatrix}x(z)\\p(z)\end{pmatrix} \qquad \text{où} \qquad p(z) = \alpha(z)\frac{\mathrm{d}x}{\mathrm{d}z}\ ,$$

et sa solution générale a donc la forme (dépendance linéaire dans les C.I.) :

$$\begin{pmatrix}x(z)\\p(z)\end{pmatrix} = \begin{pmatrix}a(z,z_0) & b(z,z_0)\\ c(z,z_0) & d(z,z_0)\end{pmatrix}\begin{pmatrix}x(z_0)\\p(z_0)\end{pmatrix} = \mathbf{M}(z,z_0)\,\begin{pmatrix}x(z_0)\\p(z_0)\end{pmatrix}\ .$$

$\mathbf{M}(z,z_0)$ est appelée **matrice de transfert** de z_0 à z. Elle possède les propriétés :

$$\mathbf{M}(z,z_0) = \mathbf{M}(z,z_1)\,\mathbf{M}(z_1,z_0) \qquad \text{et} \qquad \det\mathbf{M}(z,z_0) = 1\ .$$

La première est la loi de composition vue à la section 6.1.1 ; la seconde est établie ci-dessous.

Le choix de $p(z)$ au lieu de $\dfrac{\mathrm{d}x}{\mathrm{d}z}$ est important, car même si $\alpha(z)$ et $\beta(z)$ sont discontinus (très souvent $\alpha(z)$ et $\beta(z)$ sont constants par morceaux), $x(z)$ et $p(z)$ sont eux continus. On vérifie de plus que ce choix conduit à la **conservation du wronskien** $x_1p_2 - x_2p_1$ de deux solutions arbitaires (x_1, p_1) et (x_2, p_2) :

$$\frac{\mathrm{d}}{\mathrm{d}z}\big(x_1(z)\,p_2(z) - x_2(z)\,p_1(z)\big) = 0\ .$$

Cette conservation équivaut à $\det \mathbf{M}(z, z_0) = 1$ puisque (petit calcul) :

$$x_1(z)\,p_2(z) - x_2(z)\,p_1(z) = \big(x_1(z_0)\,p_2(z_0) - x_2(z_0)\,p_1(z_0)\big)\,\big(a(z,z_0)\,d(z,z_0) - b(z,z_0)\,c(z,z_0)\big)\ .$$

Cas importants. Lorsque α et β sont constants, on a $\dfrac{\mathrm{d}^2x}{\mathrm{d}z^2} = -\dfrac{\beta}{\alpha}\,x$, ce qui permet d'exprimer facilement $x(z)$ puis $p(z)$ en fonction de $x(z_0)$ et de $p(z_0)$. On obtient les expressions :

$$\mathbf{M}(z, z_0) = \begin{pmatrix} \cos k(z-z_0) & (\alpha k)^{-1}\,\sin k(z-z_0) \\ -\alpha k\,\sin k(z-z_0) & \cos k(z-z_0) \end{pmatrix} \quad \text{si } \frac{\beta}{\alpha} = k^2 > 0\ , \qquad (1)$$

$$\mathbf{M}(z, z_0) = \begin{pmatrix} \cosh K(z-z_0) & (\alpha K)^{-1}\,\sinh K(z-z_0) \\ \alpha K\,\sinh K(z-z_0) & \cosh K(z-z_0) \end{pmatrix} \quad \text{si } \frac{\beta}{\alpha} = -K^2 < 0\ . \qquad (2)$$

Lorsque $z \to z_0$, $\beta \to \infty$, $\beta(z - z_0) \to A$ constante (c.a.d. $\beta(z) = A\,\delta(z - z_0)$) la matrice de transfert s'écrit :

$$\mathbf{M}(z_0^+, z_0^-) = \begin{pmatrix} 1 & 0 \\ -A & 1 \end{pmatrix}\ . \qquad (3)$$

Dans ce cas singulier, $x(z)$ est continu mais $p(z)$ a une discontinuité $-Ax(z_0)$ en z_0.

Applications aux rayons. Le formalisme de la matrice de transfert est identique à celui introduit à la section 4.3.2 ; $x(z)$ et $p(z)$, produit de l'indice $n(0, 0, z)$ sur l'axe par la pente du rayon, sont les coordonnées d'un rayon. Les résultats vus alors peuvent être repris. Par exemple la condition de **conjugaison** de deux plans z_0 et z_0' s'écrit $b(z_0', z_0) = 0$, et on a $\overline{O'M_0'} = a(z_0', z_0)\,\overline{OM_0}$ où $a(z_0', z_0)$ est le grandissement (figure 21).

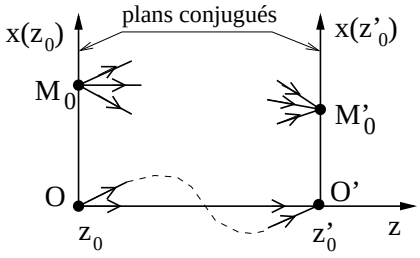


FIGURE 21

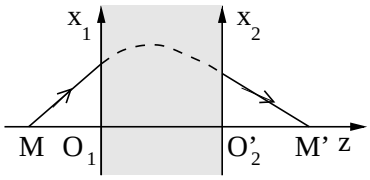


FIGURE 22

De même si la matrice (ses éléments a, b, c, d) est connue pour les deux plans d'entrée O_1 et de sortie O_2' (dans l'air) d'un système optique, la relation de conjugaison de deux points objet M et image M' situés sur l'axe (figure 22) s'écrit comme à la section 4.3.2 :

$$q' = \frac{aq + b}{cq + d} \qquad (\overline{MO_1} = q\ ,\ \overline{M'O_2'} = q')\ .$$

Enfin la propriété $\det \mathbf{M}(z, z_0) = 1$ équivaut à la conservation $\mathrm{d}x\,\mathrm{d}p = \mathrm{d}x_0\,\mathrm{d}p_0$ de l'**étendue optique** d'un pinceau lumineux (produit de ses dispersions spatiale $\mathrm{d}x$ et angulaire $\mathrm{d}p$) lors du passage d'un plan z_0 à un autre plan z.

REMARQUE. Pour des équations d'ordre n quelconque la notion de matrice de transfert persiste ainsi que la loi de composition mais la propriété $\det \mathbf{M} = 1$ n'est plus vraie en général.

6.4.2 Ondes stationnaires ; états liés ; quantification

Dans de nombreux problèmes d'ondes la solution est déterminée non par les "C.I." $\psi(z_0)$ et $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_0)$ (ou $p(z_0) = \alpha(z_0)\,\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_0)$), mais par deux **conditions aux limites (C.L.)** en deux points z_1 et $z_2 > z_1$. Ces conditions en chaque point peuvent être du type $\psi(z_i) = 0$ (tuyau sonore fermé ou corde fixée en z_i), $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_i) = 0$ (tuyau ouvert ou corde libre), ou $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_i) + \mu_i\,\psi(z_i) = 0$ (tuyau ou corde branchés sur un impédance acoustique ou mécanique, μ_i étant réel si l'impédance est imaginaire pure). Les C.L. en z_1 et z_2 entraînent automatiquement l'annulation d'une combinaison linéaire des coefficients a, b, c et d de la matrice $\mathbf{M}(z_2, z_1)$. Comme les éléments de la matrice de transfert dépendent de la pulsation ω (ou de l'énergie E en quantique), les conditions aux limites ne peuvent être satisfaites, sauf exception très rare, que pour des valeurs discrètes, en nombre fini ou infini, de ω (ou E). C'est l'origine de la **quantification des fréquences** des ondes stationnaires ou des **niveaux d'énergie**.

EXEMPLE 1. **Fréquences d'un tuyau sonore** (milieu homogène : ρ et χ constants ; longueur $L = z_2 - z_1$). On vérifie que les C.L. $\psi(z_{1,2}) = 0$, ou $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_{1,2}) = 0$, ou $\psi(z_1) = 0$ et $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_2) = 0$, s'écrivent respectivement $b = 0$, ou $c = 0$, ou $d = 0$. Ceci implique que $\mathbf{M}$ est de la forme (1). Par exemple $c = 0$ conduit à $\sin kL = 0$ et à la quantification $k_n = \dfrac{n\pi}{L} = \omega_n\sqrt{\chi\rho}$. Plus généralement, si $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(z_{1,2}) = k\,\mathrm{tg}\,\varphi_{1,2}\,\psi(z_{1,2})$, le lecteur vérifiera en procédant comme dans l'exemple 2 la condition de quantification $kL = \varphi_1 - \varphi_2 \pmod{\pi}$.

EXEMPLE 2. **Etats liés de l'E.D. de Schrödinger**. Considérons le puits de potentiel plat tel que $V(z) = V_0 < 0$ pour $0 < z < L$ et nul ailleurs, et un état lié d'énergie $E = -\dfrac{\hbar^2}{2m}K^2 < 0$. On pose $k_0^2 = \dfrac{2m}{\hbar^2}(E - V_0) > 0$. Pour les grandeurs continues ψ et $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}$ on a :

$$\begin{pmatrix} \psi(L) \\ \frac{\mathrm{d}\psi}{\mathrm{d}z}(L) \end{pmatrix} = \begin{pmatrix} \cos k_0 L & k_0^{-1}\sin k_0 L \\ -k_0 \sin k_0 L & \cos k_0 L \end{pmatrix} \begin{pmatrix} \psi(0) \\ \frac{\mathrm{d}\psi}{\mathrm{d}z}(0) \end{pmatrix} .$$

Les conditions aux limites sont $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(0) = K\,\psi(0)$ et $\dfrac{\mathrm{d}\psi}{\mathrm{d}z}(L) = -K\,\psi(L)$, puisque $\psi(z) \propto e^{Kz}$ pour $z \le 0$ et $\psi(z) \propto e^{-Kz}$ pour $z \ge L$. La quantification de l'énergie est donnée par :

$$-K = \frac{-k_o \sin k_0 L + K \cos k_0 L}{\cos k_0 L + K\,\frac{\sin k_0 L}{k_0}} \quad \text{ou} \quad k_0 L = 2\varphi \pmod{\pi} \text{ avec } \mathrm{tg}\,\varphi = \frac{K}{k_o} .$$

Cette relation devient simple pour le puits "delta" $V(z) = -\dfrac{\hbar^2}{2m}\,A\,\delta(z)$, c'est-à-dire dans la limite $|V_0| \to \infty,\ L \to 0$ et $|V_0|L = \dfrac{\hbar^2}{2m}A$ fini ($A > 0$). En effet la matrice de transfert est de la forme (3), et la condition de quantification devient $-K = K - A$; elle conduit à un seul état lié d'énergie $E = -\dfrac{\hbar^2}{2m}\,\dfrac{A^2}{4}$.
Remarque. Dans le cas d'une **barrière de potentiel** carrée de largeur L et de hauteur constante $V_0 > 0$, la matrice de transfert est de la forme (2) :

$$\mathbf{M} = \begin{pmatrix} \cosh K_0 L & K_0^{-1} \sinh K_0 L \\ K_0 \sinh K_0 L & \cosh K_0 L \end{pmatrix} \quad \text{avec} \quad K_0^2 = -\frac{2m}{\hbar^2}(E - V_0) > 0\ .$$

6.4.3 Ondes propagatives ; réflexion, transmission, adaptation d'impédance

L'étude de la réflexion et de la transmission d'ondes est un problème avec comme condition aux limites la forme de l'onde pour $z \to \infty$. Dans la suite on considère des milieux homogènes pour $z < 0$ ($\alpha(z) = \alpha_1$) et pour $z > L$ ($\alpha(z) = \alpha_2$).

▷ Réflexion et transmission

Considérons des ondes de la forme :

$$\psi(z < 0) = A_1\, e^{ik_1 z} + B_1\, e^{-ik_1 z} \quad \text{et} \quad \psi(z > L) = A_2\, e^{ik_2 z}\ .$$

Eventuellement $k_2 = iK_2$ pour une onde évanescente. Supposant connue la matrice de transfert $\mathbf{M}$ on a :

$$\begin{pmatrix} \psi(L^-) \\ \alpha \frac{\mathrm{d}\psi}{\mathrm{d}z}(L^-) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \psi(0^+) \\ \alpha \frac{\mathrm{d}\psi}{\mathrm{d}z}(0^+) \end{pmatrix}\ .$$

La continuité de ψ et de $\alpha\,\dfrac{\mathrm{d}\psi}{\mathrm{d}z}$ en 0 et L conduit aux deux équations

$$A_2\, e^{ik_2 L} = a(A_1 + B_1) + i\alpha_1 k_1 b(A_1 - B_1) \quad \text{et} \quad i\alpha_2 k_2 A_2\, e^{ik_2 L} = c(A_1 + B_1) + i\alpha_1 k_1 d(A_1 - B_1)$$

qui permettent de déterminer le coefficient de réflexion $r = \dfrac{B_1}{A_1}$ et le coefficient de transmission $t = \dfrac{A_2}{A_1}$. En physique classique ces coefficients sont souvent exprimés en fonction des **impédances des ondes** dans chaque milieu $Z_i \propto \omega^{-1}\alpha_i\, k_i$, le coefficient de proportionnalité ne dépendant pas du milieu (*cf.* exemples section 8.1.2).

Pour $L = 0$ (contact de deux milieux semi infinis), la matrice de transfert est évidemment égale à l'identité ($a = d = 1; b = c = 0$). On obtient :

$$r = \frac{\alpha_1 k_1 - \alpha_2 k_2}{\alpha_1 k_1 + \alpha_2 k_2} = \frac{Z_1 - Z_2}{Z_1 + Z_2}\ .$$

Le cas $L \neq 0$ avec un milieu intermédiaire entre 0 et L homogène ($\alpha(z) = \alpha_0,\ 0 < z < L$) est calculable, la matrice $\mathbf{M}$ étant alors de la forme (1) ou (2). Il correspond à de nombreux exemples connus en physique classique et quantique. En physique classique un cas particulier intéressant est celui de l'**adaptation d'impédance** entre les milieux 1 et 2 :

il s'agit de déterminer le milieu intermédiaire pour que le coefficient de réflexion du côté 1 soit nul, donc $B_1 = 0$. On déduit des relations ci-dessus entre A_1, B_1 et A_2 la condition :

$$i\alpha_2 k_2(a + ib\alpha_1 k_1) = c + id\alpha_1 k_1 \ .$$

Comme $\alpha_1 k_1 \neq \alpha_2 k_2$ par hypothèse, et que $a = d$, l'égalité des parties imaginaires implique $a = d = 0$. Ceci n'est possible que si $\mathbf{M}$ est de la forme (1) et $a = \cos k_0 L$ avec $k_0 L = \dfrac{\pi}{2}$ (milieu intermédiaire propagatif et de longueur L égale à un quart de longueur d'onde). Comme alors $c = -\alpha_0 k_0 = -b^{-1}$, l'égalité des parties réelles $c = -b\alpha_1 k_1 \alpha_2 k_2$ implique la condition sur les impédances $Z_0^2 = Z_1 Z_2$ (*cf.* section 8.1.2).

REMARQUE. Si les milieux 1 et 2 sont identiques (mais différents du milieu intermédiaire) on établit de même que $r = 0$ si $\sin k_0 L = 0$ soit $k_0 L = n\pi$. C'est aussi la condition de transparence du puits de potentiel plat quantique.

▷ Bilans énergétiques

Dans l'expression $\Psi(t, z) = e^{-i\omega t}\,\psi(z)$ d'une onde, la partie spatiale $\psi(z)$ est a priori complexe. Comme l'équation d'onde est à coefficients réels, $\psi(z)$ et $\overline{\psi}(z)$ constituent deux solutions auxquelles on peut appliquer la conservation du wronskien. On en déduit que la quantité réelle

$$j = \frac{i}{2}\left(\psi\,\frac{\mathrm{d}}{\mathrm{d}z}(\alpha\overline{\psi}) - \overline{\psi}\,\frac{\mathrm{d}}{\mathrm{d}z}(\alpha\psi)\right)$$

ne dépend pas de z. j est proportionnel en physique classique au **flux d'énergie** (moyenné dans le temps) et en quantique au **courant de probabilité** $i\,\dfrac{\hbar}{2m}\left(\psi\,\dfrac{\mathrm{d}\overline{\psi}}{\mathrm{d}z} - \overline{\psi}\,\dfrac{\mathrm{d}\psi}{\mathrm{d}z}\right)$ (*cf.* section 4.4.3).

Dans un milieu homogène (α et β constants) on obtient, pour des **ondes propagatives** :

$$\psi(z) = A\,e^{ikz} + B\,e^{-ikz} \quad ; \quad j = \alpha\,k(|A|^2 - |B|^2) \ ;$$

j est alors la somme algébrique des flux associés aux deux ondes se propageant en sens opposés. Pour des **ondes non propagatives** on obtient :

$$\psi(z) = A\,e^{-Kz} + B\,e^{Kz} \quad ; \quad j = 2\,\alpha K\,\Im m\,(B\overline{A}) \ ;$$

donc $j = 0$ si $B = 0$ (cas d'une onde évanescente). Le bilan énergétique pour un problème de réflexion transmission s'écrit :

$$\alpha_1 k_1(1 - |r|^2) = \alpha_2 k_2 |t|^2 \quad \text{pour } k_2 \text{ réel} \qquad \text{ou} \qquad \alpha_1 k_1(1 - |r|^2) = 0 \quad \text{pour } k_2 = iK_2 \ .$$

$|r| = 1$ si dans le milieu 2 l'onde est évanescente.

▷ Matrice S

Le cas où il n'y a qu'une seule onde $A_2\,e^{ik_2 z}$ pour $z > L$ est particulier. Dans le cas général il y en a deux : $\psi(z \geq L) = A_2\,e^{ik_2 z} + B_2\,e^{-ik_2 z}$. Si la matrice de transfert est connue, on peut exprimer A_2 et B_1, coefficients des ondes sortant de la région $[0, L]$, linéairement en fonction de A_1 et B_2, coefficients des ondes rentrant dans la région $[0, L]$. La matrice correspondante est appelée matrice $\mathbf{S}$:

$$\begin{pmatrix} A_2 \\ B_1 \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A_1 \\ B_2 \end{pmatrix} \ .$$

Elle contient la même information sur la région $[0, L]$ que la matrice de transfert. Le bilan $\alpha_1\,k_1(|A_1|^2 - |B_1|^2) = \alpha_2\,k_2(|A_2|^2 - |B_2|^2)$ montre que si les milieux 1 et 2 sont identiques, ou si on renormalise convenablement A_1, B_1 et A_2, B_2, $\mathbf{S}$ est une matrice unitaire.

6.4.4 Equations avec paramètres périodiques ; théorème de Floquet-Bloch

▷ Théorème de Floquet

Quand une E.D. linéaire a des coefficients périodiques, la **solution générale** est une somme de "modes", chacun étant le produit d'une exponentielle (réelle ou complexe) par une fonction périodique de même période que les coefficients. Lorsque la période est nulle, on retrouve le cas des coefficients constants où chaque "mode" est une simple exponentielle.

Ce résultat général se comprend bien sur l'exemple des E.D. du second ordre. Considérons le cas de l'**oscillateur paramétrique** avec une pulsation $\omega(t)$ de période T, et notons $\mathbf{M}_T$ sa matrice de transfert de $t = nT$ à $t = (n+1)T$ $(n = \cdots -1, 0, 1 \cdots)$:

$$\begin{pmatrix} x(T) \\ \dot{x}(T) \end{pmatrix} = \mathbf{M}_T \begin{pmatrix} x(0) \\ \dot{x}(0)_, \end{pmatrix} \qquad ; \qquad \mathbf{M}_T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} .$$

Si $\underline{m}$ est un vecteur propre de $\mathbf{M}_T$ correspondant à la valeur propre λ, et si $\underline{m}(t)$ désigne la solution telle que $\underline{m}(0) = \underline{m}$, alors $\underline{m}(T) = \lambda \underline{m}(0)$. Dans l'intervalle $[T, 2T]$, la solution va donc reproduire celle de l'intervalle $[0, T]$ multipliée par λ, et ainsi de suite. Ceci implique $\underline{m}(t+T) = \lambda\, \underline{m}(t)$, d'où on déduit, en posant $\underline{m}(t) = \lambda^{\frac{t}{T}} \underline{m}_{\rm per}(t) = e^{\frac{t}{T} \ln \lambda} \underline{m}_{\rm per}(t)$, que $\underline{m}_{\rm per}(t)$ est périodique ($\underline{m}_{\rm per}(t+T) = \underline{m}_{\rm per}(t)$). Il suffit alors de remarquer que des C.I. arbitraires peuvent s'écrire sous la forme de combinaisons linéaires des vecteurs propres de $\mathbf{M}_T$ (en supposant qu'ils forment une base) pour comprendre le théorème.

▷ Résonance paramétrique

Si l'un des "modes" $\underline{m}(t)$ correspond à une valeur propre de module supérieur à 1, le système est instable. Comme $\det \mathbf{M}_T = 1$, ceci se produit lorsque les deux valeurs propres λ et λ^{-1} sont réelles, c'est-à-dire pour $(a+d)^2 - 4 \geq 0$ ou $|\mathrm{Tr}\, \mathbf{M}_T| > 2$.

EXEMPLE. Considérons un oscillateur harmonique ordinaire $\ddot{x} + \omega_0^2\, x = 0$ subissant la perturbation suivante : à chaque instant nT $(n = 1, 2 \cdots)$ sa vitesse $\dot{x}(nT)$ est incrémentée de la quantité $-A\, x(nT)$ proportionnelle à son élongation à cet instant. Dans ces conditions, la matrice de transfert $\mathbf{M}'_T$ de l'oscillateur perturbé de $t = 0^+$ à $t = T^+$ est le produit de la matrice de transfert de $t = 0^+$ à $t = T^-$ de l'oscillateur harmonique non perturbé et de la matrice de transfert de T^- à T^+ qui rend compte de cette perturbation :

$$\mathbf{M}'_T = \begin{pmatrix} 1 & 0 \\ -A & 1 \end{pmatrix} \begin{pmatrix} \cos \omega_0 T & \omega_0^{-1} \sin \omega_0 T \\ -\omega_0 \sin \omega_0 T & \cos \omega_0 T \end{pmatrix} .$$

Sa trace $2 \cos \omega_0 T - \dfrac{A}{\omega_0} \sin \omega_0 T = \sqrt{4 + \dfrac{A^2}{\omega_0^2}}\, \cos(\omega_0 T + \varphi)$ est représentée en fonction

de T sur la figure 23 pour $\dfrac{A}{\omega_0} > 0$ petit ($\operatorname{tg}\varphi = \dfrac{A}{2\omega_0} \simeq \varphi$) ; les zones d'instabilité sont hachurées. La première instabilité correspond pour les perturbations à une **fréquence double** de la fréquence propre de l'oscillateur ($\omega_0 T \simeq \pi$).

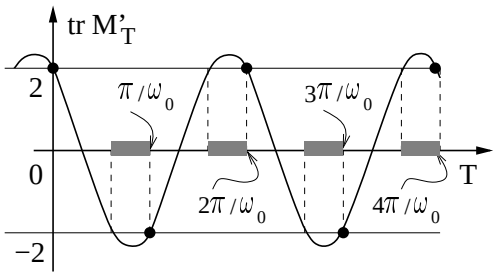


FIGURE 23

▷ Bandes d'énergie dans les cristaux

L'E.D. de Schrödinger en présence d'un potentiel périodique, de période L, a la même structure que l'équation de l'oscillateur paramétrique. Le théorème de Floquet s'appelle alors **théorème de Bloch**. Cependant les fonctions d'onde $\psi(z)$ doivent rester finies pour $z \to \pm\infty$. Ceci implique que les valeurs propres λ de la matrice de transfert $\mathbf{M}_L$ sont de module 1 et donc du type $e^{\pm ikL}$. Les facteurs exponentiels $\lambda^{z/L}$ présents dans les modes propres sont alors de la forme $e^{\pm ikz}$, c'est-à-dire des ondes planes. Cela correspond à $|\operatorname{tr}\mathbf{M}_L| = |2\cos kL| \leq 2$. On voit que à cause de la périodicité spatiale du potentiel, les niveaux d'énergie ne sont plus déterminés par des égalités entraînant leur quantification, mais par des inégalités impliquant leur regroupement par **bandes d'énergie**.

EXEMPLE. Pour une suite infinie périodique de puits de potentiel "delta ", séparés par des régions de longueur L où le potentiel est nul, la matrice de transfert $\mathbf{M}_L$ de 0^+ à L^+ est le produit de celle de la région "homogène" $V = 0$ par celle d'un "delta" (*cf.* section 6.4.2) :

$$\mathbf{M}_L = \begin{pmatrix} 1 & 0 \\ -A & 1 \end{pmatrix} \begin{pmatrix} \cos k_0 L & k_0^{-1} \sin k_0 L \\ -k_0 \sin k_0 L & \cos k_0 L \end{pmatrix} \quad \text{pour } E = \frac{\hbar^2 k_0^2}{2m} > 0$$

$$= \begin{pmatrix} 1 & 0 \\ -A & 1 \end{pmatrix} \begin{pmatrix} \cosh K_0 L & K_0^{-1} \sinh K_0 L \\ K_0 \sinh K_0 L & \cosh K_0 L \end{pmatrix} \quad \text{pour } E = -\frac{\hbar^2 K_0^2}{2m} < 0 \ .$$

Pour $E > 0$ la condition $|\operatorname{tr}\mathbf{M}_L| \leq 2$ conduit à la structure en bandes correspondant aux zones non hachurées sur la figure 23 (zones de stabilité de l'oscillateur paramétrique). Pour $E < 0$ on doit avoir $|2\cosh K_0 L - \dfrac{A}{K_0} \sinh K_0 L| < 2$, ce qui implique $A > 0$; l'état lié du puits "delta" ($K_0 = \dfrac{A}{2}$) obtenu à la section 6.4.2. est remplacé par une bande d'énergie. Inversement, en faisant tendre L vers l'infini (pour n'avoir qu'un seul "delta"), l'inégalité précédente donne $K_0 = \dfrac{A}{2}$ et on retrouve l'état lié.

6.4.5 Equations d'amplitude ; approximation adiabatique

La méthode de l'équation d'amplitude s'applique à des systèmes dont l'équation d'évolution est "proche" de celle $\ddot{x} + \omega_0^2\, x = 0$ d'un oscillateur harmonique non amorti (dont la solution générale s'écrit $x_0(t) = \dfrac{1}{2}(a_0 e^{i\omega_0 t} + \overline{a}_0 e^{-i\omega_0 t})$). Elle consiste à considérer que toute "petite" modification de cette équation se traduit sur $x_0(t)$ par un effet de modulation (d'amplitude et de phase).

La solution $x(t)$ de l'équation perturbée s'obtient en posant

$$x(t) = \frac{1}{2}(a(t)\, e^{i\omega_0 t} + \overline{a(t)}\, e^{-i\omega_0 t})$$

et en résolvant l'équation obtenue pour l'amplitude complexe $a(t)$ supposée lentement variable par rapport à $e^{\pm i\omega_0 t}$.

▷ Oscillateur paramétrique : $\ddot{x} + 2\lambda\, \dot{x} + \omega_0^2\, (1 + h\, \cos\, \gamma t)\, x = 0$

Cette équation s'écrit aussi :

$$\Big(\frac{\ddot{a}}{2} + \lambda \dot{a} + i\, \omega_0\, (\dot{a} + \lambda\, a)\Big)\, e^{i\omega_0 t} + \Big(\text{c.c.}\Big) = -\frac{1}{4}\omega_0^2\, h\, (e^{i\gamma t} + e^{-i\gamma t})(a\, e^{i\omega_0 t} + \overline{a}\, e^{-i\omega_0 t})\ .$$

A priori très compliquée, elle se simplifie beaucoup dans le cas où les paramètres $\lambda > 0$ et h sont petits devant 1, car alors $a(t)$ est une fonction lentement variable du temps. D'abord on peut négliger $\ddot{a}$ et $\lambda\, \dot{a}$ devant $\omega_0\, (\dot{a} + \lambda\, a)$; ensuite, en multipliant les deux membres par $e^{-i\omega_0 t}$, on peut ne conserver que les termes qui, comme $a(t)$, sont à variation lente. Ceci est équivalent à ne retenir dans l'équation ci-dessus que les termes dont la variation est proche de $e^{i\omega_0 t}$. On est donc amené à éliminer le terme "complexe conjugué" (qui varie essentiellement comme $e^{-i\omega_0 t}$). Quant au second membre, la seule possibilité, pour qu'il existe des termes variant comme $e^{i\omega_0 t}$, est que $\gamma = 2\omega_0 + \delta\omega$ avec $\delta\omega \ll \omega_0$. On obtient l'**E.D. d'amplitude** :

$$i\, \omega_0\, (\dot{a} + \lambda\, a) = -\frac{1}{4}\omega_0^2\, h\, e^{i\delta\omega t}\overline{a}\ .$$

Pour étudier la résonance paramétrique posons $a(t) = e^{i\frac{\delta\omega}{2}t} b(t)$. L'équation

$$\dot{b}(t) = -\Big(\lambda + i\frac{\delta\omega}{2}\Big)\, b(t) + \frac{i}{4}\omega_0\, h\, \overline{b(t)}\ ,$$

de la forme $\dot{z} = \alpha\, z + \beta\, \overline{z}$ conduit à une instabilité (*cf.* section 6.2.3) pour $\Big(\dfrac{\omega_0 h}{4}\Big)^2 \geq \lambda^2 + \Big(\dfrac{\delta\omega}{2}\Big)^2$; la zone de résonance dans le plan $(\gamma = 2\omega_0 + \delta\omega, h)$ est hachurée sur la figure 24.

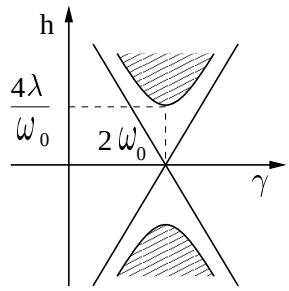


FIGURE 24

REMARQUE. Si la perturbation au lieu d'être harmonique comme $h\cos\gamma t$ est seulement périodique de période $T=\frac{2\pi}{\gamma}$, la condition de résonance $\gamma=2\omega_0$ devient $n\,\gamma=2\omega_0$; c'était le cas de l'oscillateur soumis à une suite d'impulsions considéré à la section 6.4.4.

▷ Oscillateur dont la pulsation varie adiabatiquement

Considérons l'oscillateur

$$\ddot{x}+\omega^2(t)\,x=0$$

dans le cas où la pulsation $\omega(t)$ est une fonction lentement variable du temps (temps caractéristique de variation très supérieur à la période instantanée $\dfrac{2\pi}{\omega}$). Il est alors naturel de poser :

$$x(t)=\frac{1}{2}\big(a(t)\,e^{i\int_0^t\omega(s)\,\mathrm{d}s}+\overline{a(t)}\,e^{-i\int_0^t\omega(s)\,\mathrm{d}s}\big)\ .$$

L'équation exacte $\Big(\dfrac{\ddot{a}}{2}+i\omega\dot{a}+\dfrac{i}{2}a\dot{\omega}\Big)\,e^{i\int_0^t\omega(s)\,\mathrm{d}s}+\Big(\text{complexe conjugué}\Big)=0$ conduit aussitôt à l'équation d'amplitude $\omega\,\dot{a}+\dfrac{1}{2}\,\dot{\omega}\,a=0$ et au résultat

$$\omega(t)\,|a(t)|^2=\mathrm{C}^{\mathrm{ste}}\quad\text{ ou }\quad\frac{E(t)}{\omega(t)}=\mathrm{C}^{\mathrm{ste}}\ \ (\textbf{invariant adiabatique}),$$

l'énergie E d'un oscillateur étant proportionnelle à $\omega^2|a|^2$. Donnons des exemples d'application.

- **Confinement magnétique**. On a vu (*cf.* section 6.3) que le mouvement d'une particule chargée galiléenne dans un champ magnétique constant est, pour sa composante transverse, assimilable à celui d'un oscillateur de pulsation $\left|\frac{qB}{m}\right|$ et d'énergie $\frac{1}{2}\,mv_\perp^2$. Si la composante $\vec{v}_\parallel$ est petite, et si $\overrightarrow{B}$ varie lentement d'un point de l'espace à l'autre, le mouvement transversal est celui d'un oscillateur de fréquence variant adiabatiquement ; la quantité $\frac{v_\perp^2}{B}$ est donc conservée. Comme de plus l'énergie totale $E=\frac{1}{2}m(v_\perp^2+v_\parallel^2)$ l'est aussi, le mouvement longitudinal est limité à des régions où le champ n'est pas trop grand.

- **Pression de radiation**. Une onde stationnaire pouvant être assimilée à un oscillateur, considérons une corde vibrante de longueur L fixée à ses extrémités. La condition de quantification $k=n\dfrac{\pi}{L}$ entraîne $\frac{\mathrm{d}\omega}{\omega}=-\frac{\mathrm{d}L}{L}$, d'où on déduit que lorsque la longueur L de la corde varie adiabatiquement : $\frac{\mathrm{d}E}{E}=-\frac{\mathrm{d}L}{L}$. La relation $\mathrm{d}E=-F\,\mathrm{d}L$ donne la valeur $F=\frac{E}{L}$ de la force de radiation exercée par la corde sur son extrémité. On établit de même que pour une cavité résonnante (sonore ou electromagnétique) cubique de volume $V=L^3$, la pression de radiation P, déduite de $\mathrm{d}E=-P\,\mathrm{d}V$, vaut $P=\frac{E}{3V}$ et que, pour une particule libre quantique dans une boite ($E=\hbar\omega=\frac{\hbar^2k^2}{2m}$), $P=\frac{2E}{3V}$ (pression cinétique du gaz parfait monoatomique) (*cf.* aussi section 9.4).

REMARQUE. L'E.D. de Schrödinger à une dimension étant mathématiquement semblable à l'équation de l'oscillateur, l'espace jouant le rôle du temps et $k(z)$ celui de $\omega(t)$, l'approximation adiabatique s'applique aussi en quantique lorsque la longueur caractéristique de variation du potentiel est très supérieure à la longueur d'onde locale. Une onde progressive $\Psi(z)$ est alors de la forme :

$$\Psi(z) \propto k(z)^{-\frac{1}{2}} \exp i \int^z k(z')\,\mathrm{d}z' \quad \text{avec} \quad k^2(z) = \tfrac{2m}{\hbar^2}\left(E - V(z)\right) .$$

Cette approximation connue sous le nom d'**approximation WKB** s'applique aussi aux ondes classiques.

6.5 OSCILLATEURS NON LINÉAIRES

L'étude des systèmes dynamiques non linéaires (S.D.N.L.) se justifie pour de nombreuses raisons. La première est que les lois physiques linéaires ne sont souvent que des approximations ; il faut donc au moins estimer les corrections non linéaires à apporter aux S.D.L. Une deuxième réside dans l'**instabilité** de certains S.D.L. Le fait que leur solution tende vers l'infini nécessite de revoir les hypothèses ayant conduit aux équations d'évolution, et d'examiner comment les non linéarités modifient éventuellement les conclusions de l'analyse linéaire. Enfin certains phénomènes "auto-entretenus" qui tendent asymptotiquement vers un cycle limite $x_{\lim}(t)$ ne peuvent pas découler d'une évolution linéaire, puisque $\lambda\, x_{\lim}(t)$ n'est évidemment pas solution.

Un S.D.N.L. peut bien sûr être abordé numériquement, mais son étude systématique en fonction des différents paramètres (adimensionnés) présents dans l'équation devient vite impossible et reste de toute façon "phénoménologique". Dans l'immensité du domaine non linéaire, les aspects les plus intéressants sont les changements qualitatifs de solution, appelés **bifurcations**, lorsque l'on fait varier un paramètre. Il est remarquable que leur étude repose essentiellement sur une linéarisation du problème au voisinage du point de bifurcation.

Dans la suite on utilisera principalement la méthode de l'équation d'amplitude pour étudier des exemples d'oscillateurs non linéaires. (Cette méthode constitue en fait le premier ordre d'une méthode générale : la "mise sous forme normale" d'équations près des bifurcations.) Elle s'avère appropriée en particulier pour déterminer les corrections non linéaires lorsque les coefficients des termes non linéaires sont petits.

6.5.1 Oscillateurs linéairement stables faiblement non linéaires

Considérons l'équation :

$$\ddot{x} + 2\lambda\dot{x} + \omega_0^2 x + \alpha x^2 + \beta x^3 = f\,\cos\omega t \qquad (\lambda > 0) .$$

Elle peut décrire les oscillations d'un pendule simple, auquel cas $x \equiv \theta$ est l'élongation angulaire, $\alpha = 0$ et $\beta = -\dfrac{\omega_0^2}{6}$ (si on se limite à des valeurs de θ telles que l'approximation $\sin\theta = \theta - \dfrac{\theta^3}{6}$ est valable). Elle peut aussi s'appliquer à la polarisation Nqx d'un milieu diélectrique, x décrivant les oscillations non linéaires du nuage électronique en présence d'un champ électrique oscillant ($f = \dfrac{qE}{m}$) ; les termes non linéaires sont alors dans la réalité "très petits". Souvent il apparaît un seul terme, αx^2 ou βx^3, en liaison avec l'absence ou la présence de la symétrie $x \to -x$ du problème considéré.

▷ Methode de perturbation "naïve"

L'approche la plus naturelle consiste à poser $x(t) = x_0(t) + x_1(t)$, où $x_0(t)$ est la solution de l'équation linéaire avec second membre et $x_1(t)$ une "petite" correction. En portant $x(t)$ dans l'E.D., on obtient l'E.D. linéarisée pour $x_1(t)$ (les termes en x_1^2 et x_1^3 sont négligés devant ceux en x_1) :

$$\ddot{x}_1 + 2\lambda\dot{x}_1 + \omega_0^2 x_1 + (2\alpha x_0 + 3\beta x_0^2)x_1 = -\alpha x_0^2 - \beta x_0^3 \ .$$

On peut même à priori supprimer $(2\alpha x_0 + 3\beta x_0^2)x_1$ puisque l'hypothèse $|x_1| \ll |x_0|$ entraîne par exemple $|\alpha x_0 x_1| \ll |\alpha x_0^2|$.

Cette équation, avec $x_0(t) = a_0\,\cos(\omega t + \varphi_0)$, rend compte des effets qualitatifs suivants. 1) **Optique non linéaire**. Le terme αx_0^2 (resp. βx_0^3) conduit pour $x_1(t)$, et donc pour $x(t)$, à des termes constant et de pulsation 2ω proportionnels à f^2 (resp. de pulsation ω et 3ω proportionnels à f^3). Dans l'exemple de la polarisation diélectrique, les effets associés sont connus sous le nom d'effet Kerr optique et de génération de lumière à la pulsation 2ω (resp. de dépendance de l'indice vis-à-vis de l'intensité lumineuse E^2 et de génération de lumière à la pulsation 3ω). 2) Ces harmoniques 2ω, $3\omega\cdots$ induites par les termes non linéaires d'une E.D. peuvent expliquer des effets de **résonance** de l'oscillateur pour des sous multiples $\omega = \frac{\omega_0}{2}, \frac{\omega_0}{3}\cdots$ de sa pulsation naturelle de résonance ω_0. 3) Le terme $2\alpha x_0 x_1$ peut, pour $\omega = 2\omega_0$, induire un effet de **résonance paramétrique** ; il n'est alors plus négligeable.

L'utilisation de cette méthode mérite attention. Dans le cas du mouvement libre d'un pendule $\lambda = \alpha = f = 0$, l'équation pour $x_1(t)$ avec $x_0(t) = a_0\,\cos\omega_0 t$, qui s'écrit

$$\ddot{x}_1 + \omega_0^2 x_1 = -\beta a_0^3 \cos^3\omega_0 t = -\frac{\beta a_0^3}{4}\,(3\,\cos\omega_0 t + \cos 3\omega_0 t) \ ,$$

conduit à :

$$x(t) = a_0\,\cos\omega_0 t - \frac{3\beta a_0^3}{8\omega_0}\,t\,\sin\omega_0 t + \frac{\beta a_0^3}{32\omega_0^2}\,\cos 3\omega_0 t \ .$$

La "perturbation" en $t\,\sin\omega_0 t$, due au terme "résonnant" en $\cos\omega_0 t$ présent dans le second membre de l'équation pour x_1, surprend puisqu'elle tend vers l'infini et viole la conservation de l'énergie. En fait elle n'a de sens que pout t petit ; $x(t)$ se met alors sous la forme

$$x(t) = a_0\,\cos\omega t + a_3\,\cos 3\omega t \quad \text{avec} \quad \omega = \omega_0 + \frac{3\beta a_0^2}{8\omega_0} \quad \text{et} \quad a_3 = \frac{\beta a_0^3}{32\omega_0^2} \ ,$$

et on retrouve le résultat connu que **la période dépend de l'amplitude** et qu'il y a des **harmoniques**. Une façon de montrer que ce résultat reste correct pour t grand consiste à porter a priori $x(t) = a_0\,\cos\omega t + a_3\,\cos 3\omega t$ dans l'équation du pendule. En ne gardant que les termes dominants, de pulsation ω et 3ω, on obtient

$$(\omega_0^2 - \omega^2)\,a_0\,\cos\omega t + (\omega_0^2 - 9\omega^2)\,a_3\,\cos 3\omega t + \frac{\beta a_0^3}{4}\,(3\,\cos\omega t + \cos 3\omega t) = 0 \ ,$$

d'où on déduit les mêmes expressions pour a_3 et $\omega - \omega_0$.

▷ Méthode de l'équation d'amplitude

Elle permet d'éviter le terme résonnant. Rappelons qu'elle consiste à porter

$$x(t) = \frac{1}{2}(a(t)\,e^{i\omega_0 t} + \overline{a(t)}\,e^{-i\omega_0 t})$$

dans l'équation de départ, et à ne retenir que les termes à variation proche de $e^{i\omega_0 t}$.

En remarquant que dans αx^2 il n'y a pas de tels termes, que dans βx^3 seul le terme en $a^2\overline{a}$ est de ce type, et que dans $f\cos\omega t$ il n'y a de contribution que si $\omega - \omega_0 = \delta\omega$ est petit devant ω_0, on obtient l'E.D. d'amplitude :

$$(i\omega_0 + \lambda)\,\dot{a} + i\omega_0\lambda a = \frac{1}{2}fe^{i\delta\omega t} - \frac{3}{8}\beta a^2\overline{a}\ .$$

Si $f = \lambda = 0$ (cas du **mouvement libre** d'un oscillateur non amorti anharmonique) on a : $\dot{a} = i\dfrac{3}{8}\dfrac{\beta}{\omega_0}|a|^2 a$. Cette équation, avec l'équation complexe conjuguée pour $\overline{a}$, montre que $|a|$ est constant ($\dot{a}\overline{a} + a\dot{\overline{a}} = 0$) et conduit à $a(t) = a_0 \exp\left(i\dfrac{3}{8}\dfrac{\beta}{\omega_0}|a_0|^2 t\right)$. On retrouve ainsi rapidement le résultat $\omega = \omega_0 + \dfrac{3}{8}\dfrac{\beta}{\omega_0}|a_0|^2$.

Si f et λ sont non nuls (**étude de la résonance non linéaire** près de ω_0), le régime permanent est obtenu en cherchant une solution sous la forme $a(t) = Ae^{i\varphi}e^{i\delta\omega t}$. Tenant compte de $\dfrac{\delta\omega}{\omega_0} \ll 1$ on obtient immédiatement :

$$A^2\left((\delta\omega - \frac{3}{8}\frac{\beta}{\omega_0}A^2)^2 + \lambda^2\right) = \frac{f^2}{4\omega_0^2}.$$

Les courbes $A(\delta\omega)$ se déduisent par rotation de 90^o du graphe de $\delta\omega(A)$, qui lui résulte simplement de la superposition de la courbe bien connue associée à la résonance linéaire et de la parabole $\dfrac{3}{8}\dfrac{\beta}{\omega_0}A^2$ (figure 25) :

$$\delta\omega = \frac{3}{8}\frac{\beta}{\omega_0}A^2 \pm \sqrt{\frac{f^2}{4\omega_0^2 A^2} - \lambda^2}\ .$$

On notera sur la figure 25 le déplacement de la fréquence de résonance et la constance de la valeur du maximum de l'amplitude à la résonance, et sur la figure 26, la possibilité d'hystérésis en fréquence. (En effet quand ω augmente, l'amplitude de l'oscillateur saute de B à C, et lorsque ω diminue, elle saute de D à E, car on peut montrer que les points de l'arc BD de la courbe sigmoïde correspondent à des solutions instables.)

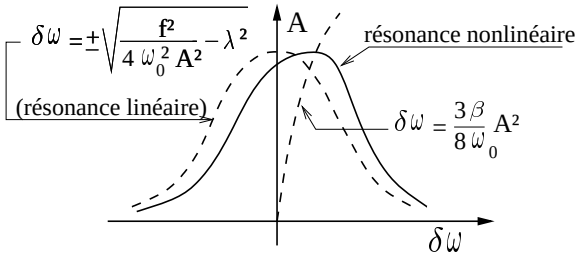


FIGURE 25

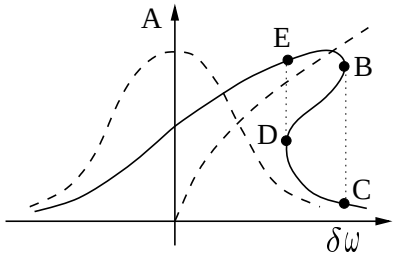


FIGURE 26

▷ Cas des oscillateurs couplés ; instabilité non linéaire

La méthode ci-dessus permet aussi d'étudier des oscillateurs décrits par les équations adimensionnées

$$\ddot{x}_1 + x_1 + \beta x_1^3 = \gamma(x_2 - x_1) \quad , \quad \ddot{x}_2 + x_2 + \beta x_2^3 = \gamma(x_1 - x_2) \ ,$$

lorsque les coefficients non linéaire β et de couplage γ sont comparables et petits. En posant $x_i(t) = \frac{1}{2}(a_i(t)\, e^{it} + \overline{a_i(t)}\, e^{-it})$, on obtient immédiatement les E.D. d'amplitudes :

$$i\,\dot{a}_1 = -\tfrac{3}{8}\beta a_1^2\overline{a_1} + \tfrac{\gamma}{2}\,(a_2 - a_1) \quad , \quad i\,\dot{a}_2 = -\tfrac{3}{8}\beta a_2^2\overline{a_2} + \tfrac{\gamma}{2}\,(a_1 - a_2) \ .$$

Elles possèdent les solutions particulières (attendues) $a_1(t) = a_2(t) = a_0 \exp\left(i\frac{3}{8}\beta|a_0|^2 t\right)$ pour le mode symétrique, et $a_1(t) = -a_2(t) = a_0 \exp\left(i\frac{3}{8}\beta|a_0|^2 t\right) \exp i\gamma t$ pour le mode antisymétrique. Un effet moins attendu, facilement vérifiable expérimentalement avec des pendules, est l'**instabilité** du mode symétrique si $\beta < 0$ et si $\frac{3}{4}|\beta||a_0|^2 > \gamma$, c.a.d. si les amplitudes d'oscillation sont assez grandes. (Le lecteur posera $a_i(t) = (a_0 + b_i(t)) \exp\left(i\frac{3}{8}\beta|a_0|^2 t\right)$, et étudiera la stabilité de l'équation linéarisée pour $z = b_1 - b_2$, qui s'écrit $i\dot{z} = -\left(\gamma + \frac{3}{8}\beta|a_0|^2\right)\, z - \frac{3}{8}\beta a_0^2\overline{z}$.) Si $\beta > 0$ c'est le mode antisymétrique qui peut devenir instable.

6.5.2 Oscillateurs linéairement instables ; exemple de Van der Pol ; bifurcations de Hopf et d'un cycle limite

L'équation adimensionnée, qui se rencontre dans de nombreux domaines de physique (mécanique, électronique...)

$$\ddot{x} - 2\epsilon(1 - x^2)\dot{x} + x = f\, \cos\omega t \quad (\epsilon > 0) \ ,$$

décrit un oscillateur (de Van der Pol) dont le coefficient d'amortissement $-2\epsilon(1 - x^2)$, négatif pour x petit, peut redevenir positif si l'amplitude est assez grande ; le terme non linéaire a donc pour effet de stabiliser le système. Sa résolution fournit un bon exemple d'application de la méthode des isoclines pour $f = 0$, et de la méthode des équations d'amplitudes pour $f \neq 0$ et ϵ petit. On discutera ensuite la "généricité" de cet exemple.

▷ Méthode des isoclines

Pour $f = 0$ l'équation s'écrit plus simplement :

$$\dot{y} + x = 0 \quad \text{avec} \quad y = \dot{x} + F(x) \quad \text{et} \quad F(x) = -2\epsilon\Big(x - \frac{x^3}{3}\Big).$$

Ceci suggère de prendre x et y, plutôt que x et $\dot{x}$, comme coordonnées dans l'espace de phase. Des relations $\mathrm{d}y + x\,\mathrm{d}t = 0$ et $\big(y - F(x)\big)\,\mathrm{d}t = \mathrm{d}x$, on déduit (en éliminant $\mathrm{d}t$) l'équation $\big(y - F(x)\big)\,\dfrac{\mathrm{d}y}{\mathrm{d}x} + x = 0$ satisfaite par les trajectoires $y(x)$, et l'équation des isoclines ($\dfrac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{tg}\,\theta$) :

$$y = -\mathrm{cotg}\,\theta\, x + F(x) \ .$$

Les isoclines $\theta = 0$ et $\theta = \dfrac{\pi}{2}$ sont respectivement l'axe Oy ($x = 0$) et la cubique $y = F(x)$; les isoclines $\theta = \pm\dfrac{\pi}{4}$ sont les cubiques $y = (-2\epsilon \mp 1)x + 2\epsilon\dfrac{x^3}{3}$.

Dans le cas $\epsilon \ll 1$ (figure 27), le lecteur se convaincra que la forme des isoclines implique qu'une trajectoire initialement proche de O s'en éloigne, tandis qu'une trajectoire initialement loin de O s'en rapproche. Ceci conduit à penser que toute trajectoire tend vers un **cycle limite**. On estime sa "taille" a en posant $x_{\lim}(t) = a\,\cos t$, et en écrivant que sur la durée d'un cycle la force de frottement ne travaille pas, et donc que sa puissance moyenne $< -2\epsilon(1-x^2)\dot{x}^2 >$ est nulle ; on obtient $a = 2$. Ce résultat est confirmé ci-dessous.

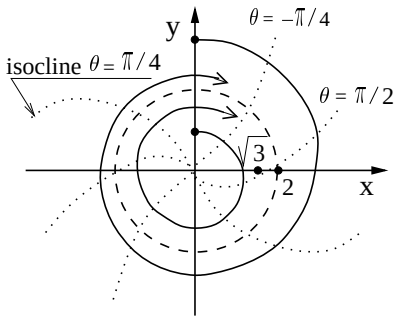


FIGURE 27

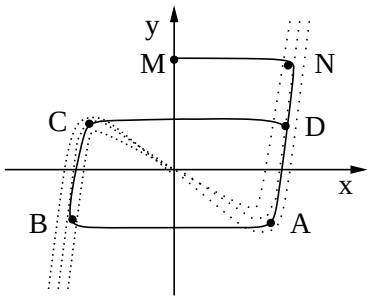


FIGURE 28

Dans le cas $\epsilon \gg 1$ (schématisé sur la figure 28), c'est le terme $-2\epsilon\Big(x - \dfrac{x^3}{3}\Big)$ qui est dominant dans l'équation des isoclines, sauf dans le voisinage immédiat de $\theta = 0$, où c'est au contraire le terme $-\mathrm{cotg}\,\theta\, x$ qui domine. Il en résulte que, à part celles qui correspondent à $\theta \simeq 0$, toutes les isoclines sont très proches de l'isocline $\theta = \dfrac{\pi}{2}$. Dès lors une trajectoire qui part d'un point M de Oy (où $\theta = 0$) reste quasiment horizontale jusqu'à ce qu'elle s'approche de N sur la courbe $y = F(X)$. Elle se trouve alors "piégée" par le réseau très serré d'isoclines jusqu'en A. Entre A et B elle redevient horizontale, puis "piégée" entre B et C, horizontale à nouveau entre C et D, pour finalement décrire indéfiniment le cycle limite $ABCDA$. On peut déduire de cette rapide analyse du portrait de phase l'allure de la fonction $x(t)$. Comme $\dot{x} = y - F(x)$, $\dot{x}$ sera petit sur les parties DA et BC du cycle et grand sur CD et AB. $x(t)$ va donc être périodique et présenter des oscillations avec des phases de variations lentes puis rapides bien séparées : ce sont des **oscillations de relaxation** (figure 29).

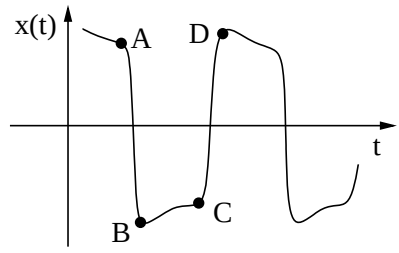


FIGURE 29

▷ Méthode de l'E.D. d'amplitude

Pour $\epsilon \ll 1$ on porte l'expression $x(t) = \frac{1}{2}\Big(a(t)\,e^{it} + \overline{a(t)}\,e^{-it}\Big)$ dans l'équation et on conserve seulement les termes à variation proche de e^{it} :

$$\dot{a} = -\frac{i}{2}\,f\,e^{i\delta\omega t} + \epsilon\,a\left(1 - \frac{|a|^2}{4}\right) \quad \text{avec} \quad \omega - 1 = \delta\omega \ll 1\ .$$

Pour $f = 0$, en posant $a(t) = A(t)e^{i\varphi(t)}$, on vérifie que φ est constant et que $\dot{A} = \epsilon\,A\Big(1 - \frac{1}{4}A^2\Big)$. Comme $\dot{A}$ est du signe de $\Big(1 - \frac{1}{4}A^2\Big)$, $A(t)$ tend toujours vers 2 et on établit ainsi l'existence du cycle limite.

Pour $f \neq 0$, toute solution stationnaire stable de la forme $a(t) = a_0\,e^{i\delta\omega t}$, pour laquelle l'oscillateur a (comme dans le cas linéaire) la même pulsation $1 + \delta\omega = \omega$ que la force, bien que sa fréquence "naturelle" soit égale à 1, caractérise ce qu'on appelle un **accrochage de fréquence**. On montre que cet accrochage a lieu si $|f| > 4|\delta\omega|$, c.a.d. si la force est assez grande. Démonstration : $|a_0|^2$ vérifie l'équation $\big((\delta\omega)^2 + \epsilon^2(1 - \frac{|a_0|^2}{4})^2\big)|a_0|^2 = \frac{1}{4}\,|f|^2$. Pour f petit elle admet une solution proche de $|a_0|^2 = 0$ et aussi, à condition que $|f| > 4|\delta\omega|$, deux solutions $|a_0|^2_\pm \simeq 4 \pm \epsilon^{-1}\sqrt{|f|^2 - 16(\delta\omega)^2}$ (proches de 4). On étudie leur stabilité en posant $a(t) = \big(a_0 + a_1(t)\big)e^{i\delta\omega t}$, et en regardant si la perturbation $a_1(t)$ tend vers zéro. L'équation linéarisée $\dot{a}_1 = \Big(\epsilon\,\big(1 - \frac{|a_0|^2}{2}\big) - i\,\delta\omega\Big)\,a_1 - \frac{1}{4}\epsilon\,a_0^2\,\overline{a_1}$ conduit aux conditions de stabilité (*cf.* section 3.2.4) $|a_0|^2 > 2$ et $\epsilon^2(1 - \frac{|a_0|^2}{2})^2 + (\delta\omega)^2 > \frac{1}{16}\epsilon^2|a_0|^4$. On vérifie que la solution $|a_0|^2_+$ est stable.

▷ Bifurcation de Hopf

L'oscillateur de Van der Pol donne un exemple de cycle limite, mais ne permet pas de décrire comment, quand on fait varier des paramètres, un oscillateur linéairement stable peut bifurquer vers un tel cycle. Près de la bifurcation la méthode de l'équation d'amplitude, qui ne garde que les termes en $e^{i\omega_0 t}$ (du type $\big(a\,e^{i\omega_0 t}\big)^p\,\big(\overline{a}\,e^{-i\omega_0 t}\big)^{(p-1)}$), conduit en l'absence de forces extérieures à l'équation générale

$$\dot{a} = \epsilon a - (\gamma + i\delta)|a|^2 a\ ,$$

limitée ci-dessus à l'ordre trois. (Van der Pol correspond à $\gamma > 0$ et $\delta = 0$, le pendule amorti à $\epsilon = -\lambda$, $\gamma = 0$ et $\delta \propto \beta$.) De l'équation $\dot{\rho} = 2\rho\,(\epsilon - \gamma\rho)$ vérifiée par $\rho = |a|^2$, on déduit en étudiant le signe de $\dot{\rho}$ que, lorsque ϵ varie, la solution $a = 0$ qui est stable pour $\epsilon < 0$ bifurque pour $\epsilon > 0$ vers le cycle limite d'amplitude $|a| = \sqrt{\dfrac{\epsilon}{\gamma}}$ si $\gamma > 0$; $-\delta\dfrac{\epsilon}{\gamma}$ est la correction non linéaire à la pulsation. Si $\gamma < 0$ on ne peut pas donner de résultat général à cet ordre du développement en puissances de a.

▷ Bifurcations d'un cycle limite

Le cycle de Van der Pol est toujours stable, mais c'est un cas particulier. Pour étudier de manière générale la stabilité d'un cycle limite $\underline{x}_{\rm lim}(t)$ de période T dans l'espace de phase (a priori à n dimensions), on pose $\underline{x}(t) = \underline{x}_{\rm lim}(t) + \underline{x}_1(t)$. L'équation linéarisée pour $\underline{x}_1(t)$ est à coefficients périodiques, et on est amené (*cf.* section 6.4.4.) à étudier sa matrice de

transfert $\mathbf{M}_T$ définie par $\underline{x}_1(T) = \mathbf{M}_T\, \underline{x}_1(0)$. Le cycle est stable si ses valeurs propres λ_i vérifient $|\lambda_i| < 1$. Quand on fait varier des paramètres, il peut devenir instable dans les trois cas suivants (de traversée du cercle unité par les λ_i ; figure 30).

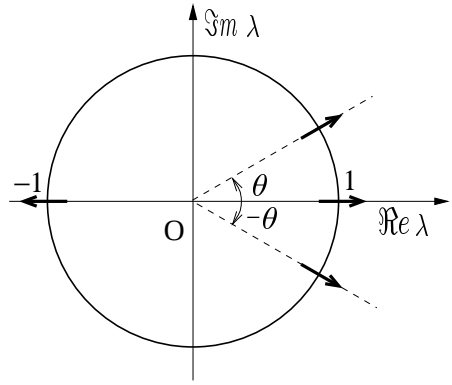


FIGURE 30

1) Une valeur propre devient égale à -1 ; la période passe brutalement de T à $2T$: **doublement de période** (figure 31).
2) Une paire de valeurs propres devient égale à $e^{\pm i\theta}$; il apparaît une nouvelle période $2\pi\,\dfrac{T}{\theta}$ et $\underline{x}(t)$ devient **quasi-périodique** (de fréquences $\dfrac{p}{T} + q\dfrac{\theta}{T}$; figure 32).
3) Une valeur propre devient égale à 1 ; il n'y pas alors de conclusion générale. $\underline{x}(t)$ peut même devenir chaotique si $n \geq 3$; on parle alors de **chaos déterministe** (voir ouvrages spécialisés).

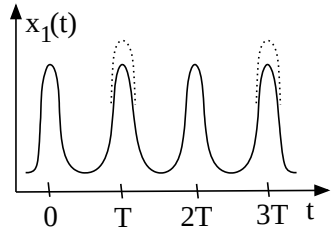


FIGURE 31

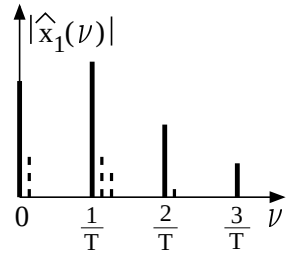


FIGURE 32

Chapitre 7

Fonctions de plusieurs variables ; analyse vectorielle

Les fonctions de plusieurs variables apparaissent dès qu'on fait de la géométrie dans l'espace (par exemple pour décrire en optique des surfaces d'onde et les rayons qui leurs sont orthogonaux), ou qu'on considère des grandeurs (champs) dépendant de l'espace et du temps. Les notions de gradient (de champs scalaires), de divergence ou rotationnel (de champs vectoriels), de dérivée dans une direction ou en suivant le mouvement, *etc.* sont des outils naturels d'analyse "locale" des champs. Ils trouvent de nombreuses applications en particulier en hydrodynamique. Cette description différentielle est reliée par les théorèmes de Stokes et d'Ostrogradski à une description intégrale basée sur les notions plus intuitives (déjà vues au chapitre 3) de circulation et de flux. Cette approche "globale" plus générale (prise en compte des discontinuités), et moins technique, est bien adaptée aux bilans de grandeurs et à une introduction à l'électromagnétisme.

Un exemple classique qui concerne des variables non spatio-temporelles est la thermodynamique des états d'équilibre où l'étude expérimentale des formes différentielles "chaleur" δQ et travail δW a joué un rôle historique pour la découverte des grandeurs énergie interne U et entropie S (premier et second principes). Même si l'approche moderne a radicalement changé (*cf.* section 1.1.4), le calcul différentiel reste un outil nécessaire, notamment pour déterminer les états d'équilibre (extremum de "potentiels") et maîtriser les changements entre les nombreux choix possibles de variables indépendantes.

7.1 CALCUL DIFFÉRENTIEL

7.1.1 Développement de Taylor ; différentielles ; variation seconde ; extremum ; E.D.P. simples

▷ Dérivées partielles

Une fonction de plusieurs variables $f(x_1, x_2 \cdots)$ (ou $f(\underline{x})$) devient une fonction d'une seule variable, x_1 par exemple, lorsque les autres $x_2, x_3 \cdots$ restent constantes. La dérivée correspondante $\dfrac{\partial f}{\partial x_1}(\underline{x})$ (ou $\partial_1 f(\underline{x})$) mesure le "taux de variation de $f(\underline{x})$ dans la direction

"x_1" au point $\underline{x}$; elle dépend à priori de $x_1, x_2 \cdots x_n$. Par exemple :

$$\partial_i x_j = \delta_{ij} \quad , \quad \partial_i \big(x_1^2 + x_2^2 + \cdots + x_n^2\big)^{\frac{1}{2}} = x_i \big(x_1^2 + x_2^2 + \cdots + x_n^2\big)^{-\frac{1}{2}} \ .$$

Si $\partial_1 f = 0$ pour tout $\underline{x}$, f ne dépend pas de x_1.

Les opérations de dérivation $\partial_1, \partial_2 \cdots \partial_n$ sur $f(\underline{x})$ peuvent être composées (comme pour une variable). Elles ont la propriété importante de commuter :

$$\frac{\partial}{\partial x_i}\,\frac{\partial}{\partial x_j} = \frac{\partial}{\partial x_j}\,\frac{\partial}{\partial x_i} \qquad \left(\text{donc } \frac{\partial^2 f}{\partial x_i\,\partial x_j} = \frac{\partial^2 f}{\partial x_j\,\partial x_i} : \textbf{ égalité de Schwarz}\right) \ .$$

On en déduit, en intégrant successivement sur x_1 et x_2, que la solution de l'équation aux dérivées partielles $\partial_2\,\partial_1 f = 0$ s'écrit : $f(\underline{x}) = f_1(x_2, x_3 \cdots x_n) + f_2(x_1, x_3 \cdots x_n)$.

La commutation ci-dessus est une conséquence de celle des opérations de translation sur les différentes variables $x_1 \to x_1 + a_1 \cdots x_n \to x_n + a_n$ et du lien (déjà vu à la section 1.1.3.) entre translation et dérivation ; par exemple :

$$f(x_1 + a_1, x_2 \cdots x_n) = \left(1 + a_1\,\partial_1 + \frac{a_1^2}{2}\,(\partial_1)^2 + \cdots\right) f(\underline{x}) = \left(e^{a_1\,\partial_1}\right)\ f(\underline{x}) \ .$$

Si on translate toutes les variables $\underline{x} \to \underline{x} + \underline{a}$, on obtient (pour des "bonnes" fonctions) le **développement de Taylor** en puissances des a_i de $\left(e^{a_1\,\partial_1}\right) \cdots \left(e^{a_n\,\partial_n}\right)\ f(\underline{x})$:

$$f(\underline{x} + \underline{a}) = f(\underline{x}) + \sum_{i=1}^{n} a_i\,\frac{\partial f}{\partial x_i}(\underline{x}) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{2}\,a_i a_j\,\frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}) + \cdots \ .$$

▷ Différentielle

La différentielle de $f(\underline{x})$ est définie de façon formelle par :

$$\mathrm{d}f(\underline{x}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x})\,\mathrm{d}x_i \ .$$

Si les $\mathrm{d}x_i$ représentent des variations "petites" des x_i, la quantité $\mathrm{d}f(\underline{x})$, "**variation infinitésimale**" de $f(\underline{x})$, est l'approximation linéaire (ordre 1) de la variation exacte $f(\underline{x} + \mathrm{d}\underline{x}) - f(\underline{x})$. Elle est très utile pour exprimer la dérivée par rapport à λ de la fonction d'une variable $f\big(\underline{x}(\lambda)\big)$ obtenue lorsque les x_i dépendent d'un paramètre λ (chemin paramétré) :

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\,f\big(\underline{x}(\lambda)\big) = \sum_i \dot{x}_i(\lambda)\,\frac{\partial f}{\partial x_i}\big(\underline{x}(\lambda)\big) \quad \text{ avec } \dot{x}_i(\lambda) = \frac{\mathrm{d}x_i(\lambda)}{\mathrm{d}\lambda} \ .$$

Pour un chemin γ allant de $\underline{x}_1 = \underline{x}(\lambda_1)$ à $\underline{x}_2 = \underline{x}(\lambda_2)$ on en déduit

$$\int_\gamma \mathrm{d}f \overset{\text{def}}{=} \int_{\lambda_1}^{\lambda_2} \frac{\mathrm{d}}{\mathrm{d}\lambda}\,f\big(\underline{x}(\lambda)\big)\,\mathrm{d}\lambda = f(\underline{x}_2) - f(\underline{x}_1) \ ,$$

résultat indépendant du chemin γ (si f est bien définie *cf.* section 7.2.2).

▷ Variation seconde

L'écart entre la variation exacte de f et sa différentielle $[f(\underline{x}+\mathrm{d}\underline{x})-f(\underline{x})]-\mathrm{d}f(\underline{x})$ est donné, si les $\mathrm{d}x_i$ sont "petits", par le deuxième terme du développement de Taylor :

$$\delta^2 f = \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i\,\partial x_j}\,\mathrm{d}x_i\,\mathrm{d}x_j = \frac{1}{2}\sum_i \mathrm{d}\left(\frac{\partial f}{\partial x_i}\right)\mathrm{d}x_i\ .$$

Interprétation géométrique (figure 1) : soient $z=\psi(x,y)$ l'équation d'une surface et, sur cette surface, un point fixe $m_0(x_0,y_0,z_0)$ et un point courant $m(x,y,z)$ proche de m_0 ($x=x_0+\mathrm{d}x$, $y=y_0+\mathrm{d}y$). Le développement de $z-z_0$ limité au premier terme

$$z-z_0=(x-x_0)\,\frac{\partial\psi}{\partial x}(x_0,y_0)+(y-y_0)\,\frac{\partial\psi}{\partial y}(x_0,y_0)$$

donne l'équation (exacte) du **plan tangent** P à la surface en m_0. La variation seconde

$$\delta^2 z=\frac{1}{2}\left(\frac{\partial^2\psi}{\partial x^2}(x_0,y_0)\,\mathrm{d}x^2+2\frac{\partial^2\psi}{\partial x\,\partial y}(x_0,y_0)\,\mathrm{d}x\,\mathrm{d}y+\frac{\partial^2\psi}{\partial y^2}(x_0,y_0)\,\mathrm{d}y^2\right)$$

représente l'écart "vertical" du point m à ce plan. Si $\delta^2 z>0$ quel que soit m proche de m_0, la surface est située au dessus du plan tangent tandis que, si $\delta^2 z<0$, elle est située en dessous (**surface localement concave**) ; si le signe de $\delta^2 z$ n'est pas défini, la surface coupe en général le plan selon deux directions obtenues en écrivant $\delta^2 z=0$ (figure 2).

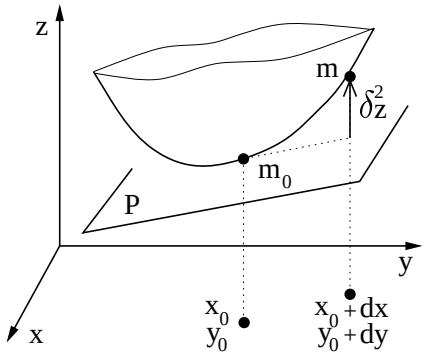


FIGURE 1

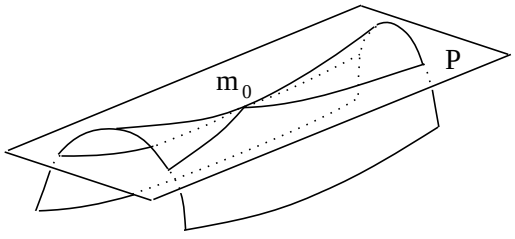


FIGURE 2

▷ Extremum

Par définition $f(\underline{x})$ est extremum (stationnaire) en $\underline{x}_0$ si :

$$\partial_i f(\underline{x}_0)=0 \quad (i=1,2\cdots n) \quad \text{ou} \quad \mathrm{d}f=0 \text{ en } \underline{x}_0\ .$$

Il s'agit d'un maximum si $\delta^2 f<0$ et d'un minimum si $\delta^2 f>0$. (Pour la surface $z=\psi(x,y)$ un extremum de ψ en m_0 correspond à un plan tangent "horizontal" en ce point.)

Extremum avec contraintes. Souvent on doit chercher un extremum de $f(\underline{x})$ lorsque les variables sont liées par des relations $C_\alpha(\underline{x})=0$ ($\alpha=1,2\cdots p<n$) qui ne permettent plus d'expliciter f comme fonction de variables indépendantes. On montre que la solution

de $\mathrm{d}f(\underline{x}) = 0$ avec $C_\alpha(\underline{x}) = 0$ s'obtient en introduisant p paramètres λ_α (**multiplicateurs de Lagrange**) et en résolvant, pour les λ_α et les x_i, les $p+n$ équations :

$$C_\alpha(\underline{x}) = 0 \quad \text{et} \quad \mathrm{d}\Big(f(\underline{x}) + \sum_\alpha \lambda_\alpha\, C_\alpha(\underline{x})\Big) = 0 \iff \frac{\partial f}{\partial x_i} = -\sum_\alpha \lambda_\alpha\, \frac{\partial C_\alpha}{\partial x_i}\ .$$

Vérification : avec cette expression pour $\partial_i f$, on a bien $\mathrm{d}f = \sum_i \partial_i f\, \mathrm{d}x_i = -\sum_\alpha \lambda_\alpha\, \mathrm{d}C_\alpha = 0$ si les contraintes sont satisfaites. (Voir applications aux sections 7.3.2 et 7.4.1,2.)

▷ E.D.P. $\sum_i a_i(\underline{x})\,\partial_i f = G(f)$ et $\sum_i a_i(f)\,\partial_i f = 0$

La résolution de ces équations aux dérivées partielles se ramène simplement à celle de **systèmes dynamiques**. En effet dans le premier cas, soient $\underline{x}(\lambda)$ les chemins définis par $\dot{\underline{x}}(\lambda) = \underline{a}\big(\underline{x}(\lambda)\big)$ et une condition initiale (C.I.) ; alors sur chacun d'eux $f\big(\underline{x}(\lambda)\big)$ obéit à l'équation différentielle ordinaire $\dfrac{\mathrm{d}f}{\mathrm{d}\lambda} = G(f)$. Dans le second cas, si on se donne la valeur $f(\underline{x}_0) = K$ de la fonction en un point, on a aussi $f(\underline{x}) = K$ sur la droite $\dot{\underline{x}}(\lambda) = \underline{a}(K)$ passant par $\underline{x}_0$ (car sur cette droite $\dfrac{\mathrm{d}f}{\mathrm{d}\lambda} = 0$).

EXEMPLE 1. **E.D.P. d'Euler** : $\displaystyle\sum_i x_i\,\partial_i f = nf$. Sur les chemins $\underline{x}(\lambda) = e^\lambda\, \underline{x}_0$ ($\dot{\underline{x}} = \underline{x}$), f vérifie $\dfrac{\mathrm{d}}{\mathrm{d}\lambda} \ln f\big(\underline{x}(\lambda)\big) = n$, d'où $f\big(e^\lambda\, \underline{x}_0\big) = e^{n\lambda}\, f(\underline{x}_0)$ pour tout $\underline{x}_0$. En posant $k = e^\lambda$ (et en changeant $\underline{x}_0$ en $\underline{x}$) on obtient $f(k\underline{x}) = k^n\, f(\underline{x})$ (**fonction homogène de degré n**). Inversement $\dfrac{\mathrm{d}}{\mathrm{d}k} f(k\underline{x}) = nk^{n-1}\, f(\underline{x})$ entraîne l'E.D.P. d'Euler pour $k = 1$. Remarque : comme $\partial_j\big[f(k\underline{x})\big] = k\big[\partial_j f\big](k\underline{x})$, $\partial_j f$ est homogène de degré $n-1$, $\partial_j\partial_k f$ de degré $n-2$, *etc.*

EXEMPLE 2. **E.D.P. d'onde simple** : $\partial_t f + c(f)\,\partial_x f = 0$. Ici $\underline{x} = (t, x)$ et $f(x,t) = K$ sur la droite $t = \lambda\,,\ x = x_0 + c(K)\,\lambda$. Par conséquent on obtient le graphe de $f(x,t)$ (en fonction de x) en translatant horizontalement chaque point $M(x)$ du graphe de $f(x,0)$ de la quantité $c\big(f(x,0)\big)\,t$. Cette E.D.P. (établie à la section 7.2.4) est la plus simple qui décrit une propagation sans déformation si la vitesse c est constante ; si c croit avec f, il y a génération d'une **onde de choc** (lorsque $f(x,t)$ n'est plus définie univoquement ; figure 3).

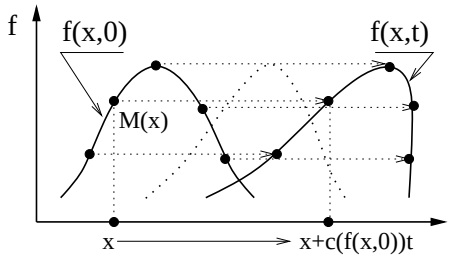


FIGURE 3

7.1.2 Dérivées spatiales de champs scalaires et vectoriels

▷ **Gradient et surfaces** (figure 4)

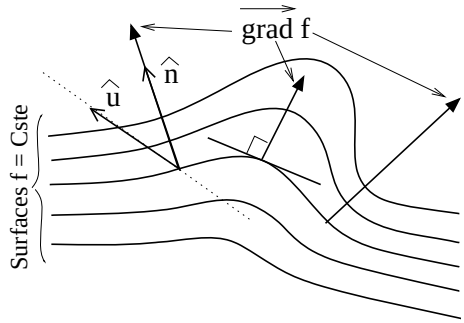


FIGURE 4

La différentielle $\mathrm{d}f = \partial_x f\,\mathrm{d}x + \partial_y f\,\mathrm{d}y + \partial_z f\,\mathrm{d}z$ d'une fonction $f(\vec{r})$ s'écrit vectoriellement :

$$\mathrm{d}f = \overrightarrow{\mathrm{grad}} f \cdot \mathrm{d}\vec{r} \quad \text{avec} \quad \overrightarrow{\mathrm{grad}} f = \frac{\partial f}{\partial x}\,\hat{x} + \frac{\partial f}{\partial y}\,\hat{y} + \frac{\partial f}{\partial z}\,\hat{z}\ .$$

Donc le champ de vecteurs $\overrightarrow{\mathrm{grad}} f$ est orthogonal aux **surfaces de niveau** $f(\vec{r}) = \mathrm{C}^{\mathrm{ste}}$ ou $\mathrm{d}f = 0$ (équipotentielles, surfaces d'onde...), et le plan tangent en $\vec{r}_0$ à une telle surface a pour équation $\overrightarrow{\mathrm{grad}} f(\vec{r}_0) \cdot (\vec{r} - \vec{r}_0) = 0$. La relation $\overrightarrow{\mathrm{grad}} f = \left|\overrightarrow{\mathrm{grad}} f\right| \hat{n}$ définit le vecteur unitaire normal à la surface $f(\vec{r}) = \mathrm{C}^{\mathrm{ste}}$.

EXEMPLE : $f(x,y,z) = z - \psi(x,y)$; $\overrightarrow{\mathrm{grad}} f = -\partial_x\psi\hat{x} - \partial_y\psi\hat{y} + \hat{z} = \left(1 + (\partial_x\psi)^2 + (\partial_y\psi)^2\right)^{\frac{1}{2}} \hat{n}$. Une **aire élémentaire** $\mathrm{d}\vec{S} = \mathrm{d}S\,\hat{n}$ sur la surface $z = \psi(x,y)$ est reliée à sa projection $\mathrm{d}S\,\hat{n}\cdot\hat{z} = \mathrm{d}x\,\mathrm{d}y$ sur le plan (x,y) par :

$$\mathrm{d}S = \mathrm{d}x\,\mathrm{d}y\,\sqrt{1 + \left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2} \simeq \mathrm{d}x\,\mathrm{d}y\,\left(1 + \frac{1}{2}\left(\frac{\partial\psi}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial\psi}{\partial y}\right)^2\right)$$

si $\left|\overrightarrow{\mathrm{grad}}\psi\right| \ll 1$.(Cette formule généralise l'expression d'un élément de longueur $\mathrm{d}l$ sur une courbe ; *cf.* section 5.1.1.)

▷ **Exemples de gradients**

Une fonction de **gradient constant** $\overrightarrow{\mathrm{grad}} f = \vec{A}$ s'écrit $f(\vec{r}) = \vec{A}\cdot\vec{r} + \mathrm{C}^{\mathrm{ste}}$, et a pour surfaces de niveau des plans (*cf.* section 3.2.2). Une fonction à **symétrie sphérique** $f(r)$ (surfaces de niveau $r = \sqrt{x^2+y^2+z^2} = \mathrm{C}^{\mathrm{ste}}$) a pour gradient $\overrightarrow{\mathrm{grad}} f = f'(r)\,\hat{r}$; en particulier $-\overrightarrow{\mathrm{grad}}\dfrac{1}{r} = \dfrac{\hat{r}}{r^2}$ (*cf.* champ électrique créé par une charge ponctuelle en O dont le potentiel est $\dfrac{q}{4\pi\epsilon_0 r}$). Si $\vec{p}$ est un vecteur fixe, $-\overrightarrow{\mathrm{grad}}\dfrac{\vec{p}\cdot\vec{r}}{r^3} = -\dfrac{\vec{p}}{r^3} + 3\dfrac{(\vec{p}\cdot\vec{r})\vec{r}}{r^5}$ donne (à $(4\pi\epsilon_0)^{-1}$ près) le champ électrique en $\vec{r}$ créé par un dipole $\vec{p}$ en O, ou (à $\dfrac{\mu_0}{4\pi}$ près) le

champ magnétique en $\vec{r}$ créé par une petite spire de moment magnétique $\vec{p} \equiv \vec{\mu} = i\,\overrightarrow{dS}$. Enfin,

$$\overrightarrow{\text{grad}}\left(\frac{\vec{p}_1\cdot\vec{p}_2}{r^3} - 3\frac{(\vec{p}_1\cdot\vec{r})\,(\vec{p}_2\cdot\vec{r})}{r^5}\right) = -\frac{3}{r^4}\left((\vec{p}_1\cdot\vec{p}_2)\,\hat{r} + \vec{p}_1(\vec{p}_2\cdot\hat{r}) + \vec{p}_2(\vec{p}_1\cdot\hat{r}) - 5(\vec{p}_1\cdot\hat{r})(\vec{p}_2\cdot\hat{r})\,\hat{r}\right)$$

donne (à $(4\pi\epsilon_0)^{-1}$ près) la force exercée par un dipole $\vec{p}_1$ en O sur un dipole $\vec{p}_2$ en $\vec{r}$ (force non colinéaire à $\vec{r}$; *cf.* section 7.3.1).

▷ Dérivée dans une direction ; opérateur nabla

Si $\overrightarrow{\mathrm{d}r} = \epsilon\,\hat{u}$ on a $\mathrm{d}f = \epsilon\,(u_x\,\partial_x f + u_y\,\partial_y f + u_z\,\partial_z f)$. Donc

$$\hat{u}\cdot\overrightarrow{\nabla} = u_x\,\frac{\partial}{\partial x} + u_y\,\frac{\partial}{\partial y} + u_z\,\frac{\partial}{\partial z}$$

correspond à la dérivation dans la direction $\hat{u}$ (figure 4). On appelle **nabla** l'opérateur

$$\overrightarrow{\nabla} = \hat{x}\,\frac{\partial}{\partial x} + \hat{y}\,\frac{\partial}{\partial y} + \hat{z}\,\frac{\partial}{\partial z} \qquad (\text{donc } \overrightarrow{\nabla} f = \overrightarrow{\text{grad}} f)$$

dont le caractère vectoriel sera justifié à la section 7.1.4.

▷ Opérateur laplacien ; solution de $\Delta f(r) = \text{Constante}$

Au second ordre : $f(\vec{r}+\epsilon\hat{u}) = f(\vec{r}) + \epsilon\,(\hat{u}\cdot\overrightarrow{\nabla})f(\vec{r}) + \dfrac{\epsilon^2}{2}\,(\hat{u}\cdot\overrightarrow{\nabla})(\hat{u}\cdot\overrightarrow{\nabla})f(\vec{r})$. Une moyenne sur les directions ($< u_i^2 >= \dfrac{1}{3}$; $< u_i u_j >= 0$ si $i \neq j$) conduit à $< f(\vec{r}+\epsilon\hat{u}) - f(\vec{r}) >= \dfrac{\epsilon^2}{6}\,\Delta f(\vec{r})$ avec :

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \quad ; \quad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \overrightarrow{\nabla}^2 \text{ opérateur Laplacien .}$$

Application : une fonction qui vérifie dans un domaine l'**E.D.P. de Laplace** $\Delta f = 0$ (potentiel électrostatique, gravitationnel...) est telle que sa valeur en un point est égale à sa valeur moyenne sur une petite sphère centrée en ce point ; ceci exclut tout maximum ou minimum à l'intérieur du domaine (donc pas d'équilibre stable pour une particule dans un tel potentiel).

Cas particuliers. Pour une fonction $f(r)$ à **symétrie sphérique**

$$\Delta f(r) = f''(r) + \frac{2}{r}\,f'(r) \quad \left(= C \text{ si } f(r) = A + \frac{B}{r} + \frac{Cr^2}{6},\ \text{et } = -k^2 f \text{ si } f \propto \frac{e^{\pm ikr}}{r}\right),$$

et à **symétrie cylindrique** ($r = \sqrt{x^2+y^2}$) :

$$\Delta f(r) = f''(r) + \frac{1}{r}\,f'(r) \quad \left(= C \text{ si } f(r) = A + B\,\ln r + \frac{Cr^2}{4}\right).$$

REMARQUES. 1) Pour $f(r = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2})$, on a :

$$\partial_i f = f'\,\frac{x_i}{r} \ , \ \partial_i{}^2 f = f''\,\left(\frac{x_i}{r}\right)^2 + f'\left(\frac{1}{r} - \frac{x_i^2}{r^3}\right) \text{ et } \Delta f = \sum_{i=1}^{n}\frac{\partial^2 f}{\partial x_i^2} = f'' + \frac{n-1}{r}\,f' .$$

2) Pour un champ de vecteurs $\overrightarrow{A}(\vec{r})$: $\Delta\overrightarrow{A} \stackrel{\text{def}}{=} \Delta A_x\,\hat{x} + \Delta A_y\,\hat{y} + \Delta A_z\,\hat{z}$.

▷ Dérivations spatiales d'un champ de vecteurs $\overrightarrow{A}(\vec{r})$

Les dérivées partielles relatives aux composantes sont au nombre de neuf ($\partial_i A_j$, $i,j = 1,2,3$). Compte tenu du caractère vectoriel de $\overrightarrow{\nabla}$, on peut former avec $\overrightarrow{A}$ (*cf.* section 3.2.1) un champ scalaire, **divergence** de $\overrightarrow{A}$,

$$\mathrm{div}\,\overrightarrow{A} = \overrightarrow{\nabla}\cdot\overrightarrow{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\ ,$$

un champ vectoriel, **rotationnel** de $\overrightarrow{A}$,

$$\overrightarrow{\mathrm{rot}}\,\overrightarrow{A} = \overrightarrow{\nabla}\wedge\overrightarrow{A} = (\partial_y A_z - \partial_z A_y)\,\hat{x} + (\partial_z A_x - \partial_x A_z)\,\hat{y} + (\partial_x A_y - \partial_y A_x)\,\hat{z}\ ,$$

et aussi un champ "quadrupolaire" (*cf.* section 4.3.1) de composantes :

$$\frac{1}{2}\,(\partial_i A_j + \partial_j A_i) - \frac{1}{3}\,\mathrm{div}\,\overrightarrow{A}\,\delta_{ij}\ .$$

Interprétation (figures 5a,b) : on verra à la section 7.2.3. que $\mathrm{div}\,\overrightarrow{A}\,\mathrm{d}V$ est le **flux** de $\overrightarrow{A}$ à travers la surface infinitésimale qui limite le volume $\mathrm{d}V$, et que le flux $\overrightarrow{\mathrm{rot}}\,\overrightarrow{A}\cdot\mathrm{d}\overrightarrow{S}$ est la **circulation** de $\overrightarrow{A}$ le long du circuit infinitésimal fermé frontière de $\mathrm{d}S$ (résultats importants utilisés à la section 7.1.3).

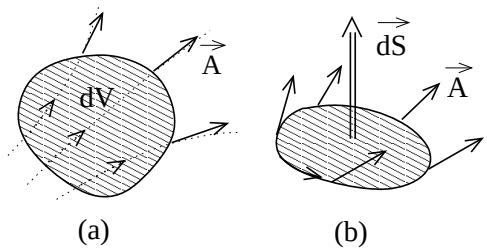


FIGURE 5

▷ Exemples de divergences et de rotationnels (Les calculs, élémentaires, sont laissés au lecteur.)

Pour un champ à **symétrie sphérique** :

$$\mathrm{div}\,f(r)\hat{r} = \frac{2}{r}\,f(r) + f'(r)\ \ (=0 \text{ si } f(r) = \frac{C}{r^2}) \quad \text{et} \quad \overrightarrow{\mathrm{rot}}\,f(r)\hat{r} = 0\ .$$

Pour un champ à **symétrie cylindrique** ($r = \sqrt{x^2+y^2}$, $\hat{r} = \left(\frac{x}{r}, \frac{y}{r}, 0\right)$ et $\hat{\theta} = \left(-\frac{y}{r}, \frac{x}{r}, 0\right)$), il y a trois cas :

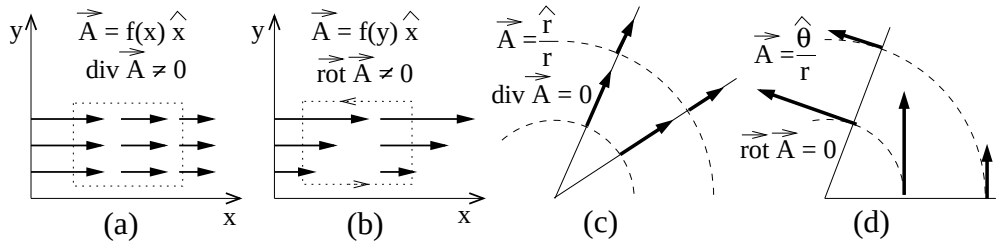
- champ selon z : $\mathrm{div}\,f(r)\,\hat{z} = 0$, $\overrightarrow{\mathrm{rot}}\,f(r)\,\hat{z} = -f'(r)\,\hat{\theta}$ $(=0 \text{ si } f(r) = C)$;

- champ radial : $\mathrm{div} f(r)\,\hat{r} = \dfrac{1}{r} f(r) + f'(r)$ $\left(= 0 \text{ si } f(r) = \dfrac{C}{r}\right)$, $\overrightarrow{\mathrm{rot}}\, f(r)\hat{r} = 0$;

- champ orthoradial : $\mathrm{div} f(r)\,\hat{\theta} = 0$, $\overrightarrow{\mathrm{rot}}\, f(r)\,\hat{\theta} = \left(\dfrac{1}{r} f(r) + f'(r)\right)\hat{z}$ $\left(= 0 \text{ si } f(r) = \dfrac{C}{r}\right)$.

Autres exemples. Si $\vec{\omega}$ est un vecteur fixe (par exemple $(0,0,\omega)$), on a $\mathrm{div}(\vec{\omega} \wedge \vec{r}) = 0$ et $\overrightarrow{\mathrm{rot}}\, (\vec{\omega} \wedge \vec{r}) = 2\,\vec{\omega}$; pour le **champ des vitesses d'un solide** $\vec{v} = \vec{v}_0 + \vec{\omega} \wedge \vec{r} \Longrightarrow \overrightarrow{\mathrm{rot}}\, \vec{v} = 2\,\vec{\omega}$; pour un **champ magnétique constant** $\overrightarrow{B} = \overrightarrow{\mathrm{rot}}\, \overrightarrow{A}$ le potentiel vecteur est $\overrightarrow{A} = \dfrac{1}{2}\, \overrightarrow{B} \wedge \vec{r}$ (à un gradient près). De même sur l'exemple $\vec{p} = (0,0,p)$ constant, on vérifie $\mathrm{div}\left(\dfrac{\vec{p} \wedge \vec{r}}{r^3}\right) = 0$ et $\overrightarrow{\mathrm{rot}}\, \left(\dfrac{\vec{p} \wedge \vec{r}}{r^3}\right) = -\dfrac{1}{r^3}\,\vec{p} + \dfrac{3(\vec{p}\cdot\vec{r})\vec{r}}{r^5}$ $\left(= -\overrightarrow{\mathrm{grad}}\, \dfrac{\vec{p}\cdot\vec{r}}{r^3}\right.$ d'où deux façons d'écrire le champ d'un dipole) ; en prenant $\overrightarrow{\mathrm{d}l} \parallel \hat{z}$, on vérifie : $\overrightarrow{\mathrm{rot}}\, \left(\dfrac{I\,\overrightarrow{\mathrm{d}l}}{r}\right) = I\overrightarrow{\mathrm{d}l} \wedge \dfrac{\hat{r}}{r^2}$ (*cf.* formule de **Biot et Savart** déduite du potentiel vecteur pour un élément de courant situé en O).

REMARQUE de vocabulaire. On prendra garde qu'un champ tel que $\mathrm{div}\,\overrightarrow{A} \neq 0$ (resp. $\overrightarrow{\mathrm{rot}}\, \overrightarrow{A} \neq 0$) n'a pas nécessairement des lignes de champ qui divergent (resp. tournent) (figures 6a,b). Inversement un champ dont les lignes divergent (resp. tournent) peut vérifier $\mathrm{div}\,\overrightarrow{A} = 0$ (resp. $\overrightarrow{\mathrm{rot}}\, \overrightarrow{A} = 0$) (*cf.* $\overrightarrow{A} = \dfrac{\hat{r}}{r}$ et $\overrightarrow{A} = \dfrac{\hat{\theta}}{r}$ en symétrie cylindrique ; figures 6c,d).



FIGURES 6

▷ Propriétés de l'opérateur nabla

$\overrightarrow{\nabla}$ possède la double propriété de se comporter comme un **opérateur de dérivation** et comme un **vecteur** dans des formules telles que celles du double produit vectoriel ou du produit mixte. Cette remarque permet d'écrire directement de nombreuses relations concernant les gradients, divergences et rotationnels, sans avoir à passer par les composantes. Les plus utilisées concernant les gradients et les divergences sont

$\overrightarrow{\nabla}(fg) = f\overrightarrow{\nabla} g + g\overrightarrow{\nabla} f \quad (\overrightarrow{\mathrm{grad}}(fg) = f\overrightarrow{\mathrm{grad}} g + g\overrightarrow{\mathrm{grad}} f)$,

$\overrightarrow{\nabla} \cdot \overrightarrow{\nabla} f = \Delta f \quad (= \mathrm{div}\overrightarrow{\mathrm{grad}} f)$,

$\overrightarrow{\nabla} \cdot (f\overrightarrow{A}) = \overrightarrow{\nabla} f \cdot \overrightarrow{A} + f\overrightarrow{\nabla} \cdot \overrightarrow{A} \quad (\mathrm{div}(f\overrightarrow{A}) = \overrightarrow{\mathrm{grad}} f \cdot \overrightarrow{A} + f\,\mathrm{div}\overrightarrow{A})$,

$\overrightarrow{\nabla} \cdot (\overrightarrow{\nabla} \wedge \overrightarrow{A}) = 0 \quad (\mathrm{div}(\overrightarrow{\mathrm{rot}}\, \overrightarrow{A}) = 0)$,

$\overrightarrow{\nabla} \cdot (\overrightarrow{A} \wedge \overrightarrow{B}) = \overrightarrow{B} \cdot (\overrightarrow{\nabla} \wedge \overrightarrow{A}) - \overrightarrow{A} \cdot (\overrightarrow{\nabla} \wedge \overrightarrow{B}) \quad (\mathrm{div}(\overrightarrow{A} \wedge \overrightarrow{B}) = \overrightarrow{B} \cdot \overrightarrow{\mathrm{rot}}\, \overrightarrow{A} - \overrightarrow{A} \cdot \overrightarrow{\mathrm{rot}}\, \overrightarrow{B})$,

et, concernant les rotationnels :

$$\overrightarrow{\nabla} \wedge (f\overrightarrow{A}) = \overrightarrow{\nabla} f \wedge \overrightarrow{A} + f\overrightarrow{\nabla} \wedge \overrightarrow{A} \quad (\overrightarrow{\mathrm{rot}}\,(f\overrightarrow{A}) = \overrightarrow{\mathrm{grad}} f \wedge \overrightarrow{A} + f\overrightarrow{\mathrm{rot}}\,\overrightarrow{A}) \ ,$$
$$\overrightarrow{\nabla} \wedge (\overrightarrow{\nabla} f) = 0 \quad (\overrightarrow{\mathrm{rot}}\,\overrightarrow{\mathrm{grad}} f = 0) \ ,$$
$$\overrightarrow{\nabla} \wedge (\overrightarrow{\nabla} \wedge \overrightarrow{A}) = \overrightarrow{\nabla}(\overrightarrow{\nabla} \cdot \overrightarrow{A}) - \Delta\overrightarrow{A} \quad (\overrightarrow{\mathrm{rot}}\,(\overrightarrow{\mathrm{rot}}\,\overrightarrow{A}) = \overrightarrow{\mathrm{grad}}\,(\mathrm{div}\,\overrightarrow{A}) - \Delta\overrightarrow{A}) \ ,$$
$$(\overrightarrow{A} \cdot \overrightarrow{\nabla})\overrightarrow{A} = \overrightarrow{\nabla}\frac{\overrightarrow{A}^2}{2} + (\overrightarrow{\nabla} \wedge \overrightarrow{A}) \wedge \overrightarrow{A} \ ,$$
$$\overrightarrow{\nabla} \wedge (\overrightarrow{A} \wedge \overrightarrow{B}) = (\overrightarrow{\nabla} \cdot \overrightarrow{B})\overrightarrow{A} + (\overrightarrow{B} \cdot \overrightarrow{\nabla})\overrightarrow{A} - (\overrightarrow{\nabla} \cdot \overrightarrow{A})\overrightarrow{B} - (\overrightarrow{A} \cdot \overrightarrow{\nabla})\overrightarrow{B} \ .$$

DÉMONSTRATION (un exemple). Pour calculer $\overrightarrow{\nabla} \wedge (\overrightarrow{A} \wedge \overrightarrow{B})$ on écrit $\overrightarrow{\nabla} = \overrightarrow{\nabla}_A + \overrightarrow{\nabla}_B$ où par définition $\overrightarrow{\nabla}_A$ n'agit que sur $\overrightarrow{A}$ et $\overrightarrow{\nabla}_B$ que sur $\overrightarrow{B}$. Alors :

$$\overrightarrow{\nabla}_A \wedge (\overrightarrow{A} \wedge \overrightarrow{B}) = (\overrightarrow{B} \cdot \overrightarrow{\nabla}_A)\overrightarrow{A} - (\overrightarrow{\nabla}_A \cdot \overrightarrow{A})\overrightarrow{B} = (\overrightarrow{B} \cdot \overrightarrow{\nabla})\overrightarrow{A} - (\overrightarrow{\nabla} \cdot \overrightarrow{A})\overrightarrow{B} \ .$$

De même $\overrightarrow{\nabla}_B \wedge (\overrightarrow{A} \wedge \overrightarrow{B}) = (\overrightarrow{\nabla}_B \cdot \overrightarrow{B})\overrightarrow{A} - (\overrightarrow{A} \cdot \overrightarrow{\nabla}_B)\overrightarrow{B} = (\overrightarrow{\nabla} \cdot \overrightarrow{B})\overrightarrow{A} - (\overrightarrow{A} \cdot \overrightarrow{\nabla})\overrightarrow{B}$.

7.1.3 Dérivées temporelles et applications hydrodynamiques

▷ Dérivées partielle et particulaire

En physique les champs $f(\vec{r}, t)$ dépendent en général du temps. La dérivée partielle $\dfrac{\partial f}{\partial t}$ (ou $\partial_t f$) décrit l'évolution de f en un point $\vec{r}$ fixé. On peut aussi être intéressé, par exemple dans le cas d'un fluide en mouvement dont le champ de vitesse est $\vec{v}(\vec{r}, t)$, par ce que serait l'évolution locale de f "en suivant le fluide" (qui se déplace de $\epsilon\vec{v}$ dans l'intervalle de temps ϵ). La limite $\epsilon \to 0$ de $\epsilon^{-1}\,(f(\vec{r} + \epsilon\vec{v}, t + \epsilon) - f(\vec{r}, t))$ égale à

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \left(\frac{\partial}{\partial t} + \vec{v} \cdot \overrightarrow{\nabla}\right) f$$

est la **dérivée particulaire** (ou en suivant le mouvement) notée aussi en abrégé $\mathrm{d}_t f$. On remarquera que $\mathrm{d}_t f$ est fonction de $\vec{r}$ et t, et que la définition entraîne que $\mathrm{d}_t = \partial_t + \vec{v} \cdot \overrightarrow{\nabla}$ possède toutes les propriétés d'une dérivation : $\mathrm{d}_t\,(fg) = f\,\mathrm{d}_t g + g\,\mathrm{d}_t f$, $\mathrm{d}_t\,(\overrightarrow{A} \cdot \overrightarrow{B}) = (\mathrm{d}_t\overrightarrow{A}) \cdot \overrightarrow{B} + \overrightarrow{A} \cdot \mathrm{d}_t\overrightarrow{B}$... Noter aussi que la propriété $\partial_t f = 0$ caractéristique de la **stationnarité** d'un champ n'implique en général pas $\mathrm{d}_t f = 0$ (et réciproquement).

EXEMPLES. Si $f(\vec{r}, t) = x$ alors $\partial_x x = 1$, $\partial_t x = \partial_y x = \partial_z x = 0$ et $\mathrm{d}_t x = v_x(\vec{r}, t)$. Plus généralement $\mathrm{d}_t\vec{r} = \vec{v}$ est le **champ de vitesse** et $\mathrm{d}_t\vec{v} = \vec{a}$ le champ d'accélération du fluide. Si $f(\vec{r}, t) = 0$ est l'équation de la surface de séparation (évolutive) de deux fluides 1 et 2, le fait que, par définition, un fluide ne puisse pas passer de $f < 0$ à $f > 0$ conduit à $\mathrm{d}_t f = 0$ sur la surface ; ceci entraîne la **continuité de la composante normale de la vitesse** puisque :

$$\mathrm{d}_t f = \partial_t f + \vec{v}_1 \cdot \overrightarrow{\nabla} f = \partial_t f + \vec{v}_2 \cdot \overrightarrow{\nabla} f \ .$$

Approximation $\mathrm{d}_t \simeq \partial_t$. Pour des ondes sonores planes la condition "$\vec{v} \cdot \overrightarrow{\nabla} \ll \partial_t$" est satisfaite si $v \ll \dfrac{\omega}{k}$ (vitesse d'oscillation du fluide très inférieure à la vitesse du son, ou déplacement très inférieur à la longueur d'onde).

▷ Champ de vitesse et déformations (figure 7)

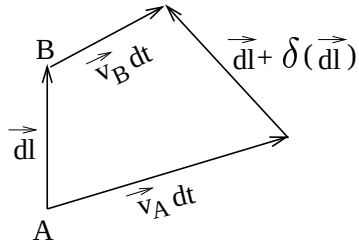


FIGURE 7

Le vecteur $\overrightarrow{AB}$ joignant deux points entraînés par le fluide varie de $\delta(\overrightarrow{AB}) = (\vec{v}_B - \vec{v}_A)\,\mathrm{d}t$ entre les instants t et $t + \mathrm{d}t$. Si A et B sont voisins ($\overrightarrow{AB} = \overrightarrow{\mathrm{d}l}$), alors $\delta(\overrightarrow{\mathrm{d}l}) = \overrightarrow{\mathrm{d}v}\,\mathrm{d}t$ ou encore :

$$\delta(\overrightarrow{\mathrm{d}l}) = \mathrm{d}t\,\left(\frac{1}{2}\,\overrightarrow{\mathrm{rot}}\,\vec{v} \wedge \overrightarrow{\mathrm{d}l} + [\dot{\epsilon}]\,\overrightarrow{\mathrm{d}l}\right) \qquad \text{avec} \quad [\dot{\epsilon}]_{ij} = \frac{1}{2}\,(\partial_i v_j + \partial_j v_i)\ .$$

DÉMONSTRATION. On vérifie pour $\delta(\mathrm{d}l_x)$ par exemple : $\frac{1}{2}\Big((\partial_z v_x - \partial_x v_z)\,\mathrm{d}z - (\partial_x v_y - \partial_y v_x)\,\mathrm{d}y\Big) + \frac{1}{2}\Big(2\partial_x v_x\,\mathrm{d}x + (\partial_x v_y + \partial_y v_x)\,\mathrm{d}y + (\partial_x v_z + \partial_z v_x)\,\mathrm{d}z\Big) = \partial_x v_x\,\mathrm{d}x + \partial_y v_x\,\mathrm{d}y + \partial_z v_x\,\mathrm{d}z = \mathrm{d}v_x$.

La première contribution à $\delta(\overrightarrow{\mathrm{d}l})$ correspond à une **rotation** de vitesse angulaire $\vec{\omega} = \frac{1}{2}\,\overrightarrow{\mathrm{rot}}\,\vec{v}$. La seconde correspond à une **déformation**, et entraîne **une variation relative de volume** $\mathrm{d}t\,\mathrm{tr}[\dot{\epsilon}] = \mathrm{d}t\,\mathrm{div}\vec{v}$ (*cf.* section 4.3.1). On en déduit que, en suivant le mouvement (*cf.* aussi section 7.2.3) :

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(\mathrm{d}^3 r) = \mathrm{div}\vec{v}\,\mathrm{d}^3 r\ .$$

Un écoulement est **irrotationnel** si $\overrightarrow{\mathrm{rot}}\,\vec{v} = 0$ et **incompressible** si $\mathrm{div}\vec{v} = 0$.

▷ Théorèmes de Bernouilli, Kelvin, Helmholtz et Lagrange

Le champ des vitesses d'un fluide possède des propriétés remarquables lorsque sa dérivée particulaire est un gradient :

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \frac{\partial \vec{v}}{\partial t} + (\vec{v}\cdot\overrightarrow{\nabla})\vec{v} = -\overrightarrow{\mathrm{grad}} f$$

(par exemple $f = \dfrac{P}{\rho} + gz$ pour un fluide incompressible dans le champ de pesanteur ; *cf.* section 7.2.4).

- **Théorème de Bernouilli** : si l'écoulement est stationnaire, alors $\dfrac{\vec{v}^{\,2}}{2} + f$ est constant le long d'une ligne de courant et, s'il est de plus irrotationnel, alors $\dfrac{\vec{v}^{\,2}}{2} + f$ est uniforme.

DÉMONSTRATION : la stationnarité implique $\mathrm{d}_t \equiv \vec{v}\cdot\overrightarrow{\nabla}$; en particulier :

$$-\mathrm{d}_t f = -\vec{v}\cdot\overrightarrow{\nabla} f = \vec{v}\cdot\mathrm{d}_t\vec{v} = \mathrm{d}_t\Big(\frac{\vec{v}^{\,2}}{2}\Big) \quad \text{soit} \quad \mathrm{d}_t\Big(\frac{\vec{v}^{\,2}}{2}+f\Big) = 0\ .$$

De plus (*cf.* $\overrightarrow{\nabla}\dfrac{\overrightarrow{A}^2}{2}$ ci-dessus) $-\overrightarrow{\nabla} f = (\vec{v}\cdot\overrightarrow{\nabla})\vec{v} = (\overrightarrow{\nabla}\wedge\vec{v})\wedge\vec{v} + \overrightarrow{\nabla}\Big(\dfrac{\vec{v}^{\,2}}{2}\Big)$ soit, dans le cas irrotationnel $\overrightarrow{\nabla}\Big(\dfrac{\vec{v}^{\,2}}{2}+f\Big) = 0$. (Ce résultat s'étend à $\overrightarrow{\nabla}\Big(\dfrac{\vec{v}^{\,2}}{2}+f+\partial_t\varphi\Big)$ si $\vec{v} = \overrightarrow{\nabla}\varphi$ n'est pas stationnaire.)

- **Théorème de Kelvin** : la circulation de $\vec{v}$ sur un contour fermé emporté par le fluide reste constante, soit

$$\mathrm{d}_t\Big(\oint \vec{v}\cdot\overrightarrow{\mathrm{d}l}\Big) = 0\ .$$

DÉMONSTRATION : $\mathrm{d}_t\big(\oint \vec{v}\cdot\overrightarrow{\mathrm{d}l}\big) = \oint(\mathrm{d}_t\vec{v})\cdot\overrightarrow{\mathrm{d}l} + \oint\vec{v}\cdot(\mathrm{d}_t\overrightarrow{\mathrm{d}l}) = -\oint\overrightarrow{\mathrm{grad}}f\cdot\overrightarrow{\mathrm{d}l} + \oint\vec{v}\cdot\mathrm{d}\vec{v} = \oint\mathrm{d}\big(-f+\frac{\vec{v}^{\,2}}{2}\big) = 0.$

- **Théorème d'Helmholtz** : les tubes de vortex (composés de lignes de vortex parallèles à $\overrightarrow{\mathrm{rot}}\ \vec{v} = 2\vec{\omega}$) sont entraînés par le fluide et conservent leur "force" (flux de $\vec{\omega}$ à travers une section du tube). *DÉMONSTRATION* (figure 8) : la surface d'un tube étant par définition tangente en tout point à $\vec{\omega}$, le flux $\vec{\omega}\cdot\overrightarrow{\mathrm{d}S}$ à travers tout élément $\overrightarrow{\mathrm{d}S}$ de cette surface est nul, et donc aussi la circulation de $\vec{v}$ "autour" de $\overrightarrow{\mathrm{d}S}$. Or cette circulation se conserve dans le transport (théorème de Kelvin). Pour établir le second résultat on choisit $\overrightarrow{\mathrm{d}S} \parallel \vec{\omega}$; $\vec{\omega}\cdot\overrightarrow{\mathrm{d}S}$ est une caractéristique du tube car $\mathrm{div}\,\overrightarrow{\omega} = 0$ (considérer un volume de tube limité par deux sections droites).

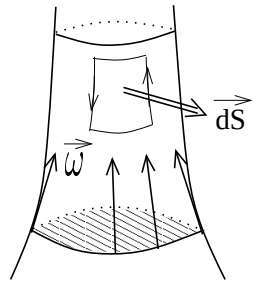


FIGURE 8

- **Théorème de Lagrange** : un écoulement initialement irrotationel ($\vec{\omega} = 0$ partout) reste irrotationnel. *DÉMONSTRATION* : si $\vec{\omega}$ devenait non nul quelque part, la circulation de $\vec{v}$ "autour" d'un élément $\mathrm{d}\vec{S}$ parallèle à $\vec{\omega}$ serait non nul, ce qui contredit le théorème de Kelvin.

Ces théorèmes sont à la base de la physique des tourbillons (*cf.* ouvrages spécialisés).

7.1.4 Changements de variables ; règles de calcul

▷ Définition et mise en garde

Si les n variables x_i sont des fonctions de p variables indépendantes x'_j, l'égalité

$$f(x_1, x_2\cdots x_n) = f\big(x_1(\underline{x}'), x_2(\underline{x}')\cdots x_n(\underline{x}')\big) = g(x'_1, x'_2\cdots x'_p)$$

définit une nouvelle fonction g des x'_j différente de f (même si $p = n$). Cependant l'habitude (fâcheuse) en physique est de conserver la même appellation f, car on a affaire à la même grandeur ; par exemple pour l'entropie, on désigne toujours par S les fonctions pourtant différentes $S(U,V)$ et $S(P,T)$. On corrige si nécessaire cette ambiguïté lors des calculs de dérivées partielles en précisant en indice les variables mises en jeu, par exemple en écrivant :

$$\frac{\partial}{\partial x'_1} f\big(x_1(\underline{x}'), x_2(\underline{x}') \cdots x_n(\underline{x}')\big) = \frac{\partial f}{\partial x'_1}\Big|_{x'_2, x'_3 \cdots x'_p} .$$

EXEMPLE. L'énergie d'un condensateur dont la capacité C dépend d'un paramètre x s'écrit :

$$E = \frac{1}{2}\, C(x) V^2 = \frac{1}{2}\, \frac{Q^2}{C(x)} \quad (Q = C(x)V) .$$

On a $\dfrac{\partial E}{\partial x}\Big|_V = \dfrac{1}{2}\, C'(x) V^2$ à ne pas confondre avec $\dfrac{\partial E}{\partial x}\Big|_Q = -\dfrac{1}{2}\, \dfrac{C'(x)}{C^2(x)}\, Q^2 = -\dfrac{1}{2}\, C'(x) V^2$ (*cf.* section 7.3.1.).

▷ Règles de calcul

La différentielle de la nouvelle fonction et ses dérivées partielles s'obtiennent en portant les différentielles $\mathrm{d}x_i = \sum\limits_j \dfrac{\partial x_i(\underline{x}')}{\partial x'_j}\, \mathrm{d}x'_j$ dans l'expression originale de $\mathrm{d}f$, et en regroupant les termes en $\mathrm{d}x'_j$:

$$\mathrm{d}f = \sum_i \frac{\partial f}{\partial x_i}\, \mathrm{d}x_i = \sum_{i,j} \frac{\partial f}{\partial x_i}\, \frac{\partial x_i}{\partial x'_j}\, \mathrm{d}x'_j = \sum_j \frac{\partial f}{\partial x'_j}\, \mathrm{d}x'_j \quad \text{avec} \quad \frac{\partial f}{\partial x'_j} = \sum_i \frac{\partial x_i}{\partial x'_j}\, \frac{\partial f}{\partial x_i} ,$$

d'où la règle pour les opérations de dérivation :

$$\frac{\partial}{\partial x'_j} = \sum_i \frac{\partial x_i}{\partial x'_j}\, \frac{\partial}{\partial x_i} .$$

EXEMPLE 1. **E.D.P. de propagation** $\left(\partial_x^2 - \dfrac{1}{c^2}\partial_t^2\right) f = \left(\partial_x + \dfrac{1}{c}\partial_t\right)\left(\partial_x - \dfrac{1}{c}\partial_t\right) f = 0$. Si $u = x + ct$ et $v = x - ct$, alors $\partial_x = \partial_u + \partial_v$ et $\partial_t = c(\partial_u - \partial_v)$. L'équation devient $4\partial_u \partial_v f = 0$ et a pour solution $f_1(x - ct) + f_2(x + ct)$ (*cf.* section 7.1.1).

EXEMPLE 2. **Caractère vectoriel de** $\overrightarrow{\nabla}$. Dans une rotation de θ du référentiel xOy, on a, pour les coordonnées

$$x = x' \cos\theta - y' \sin\theta \quad \text{et} \quad y = x' \sin\theta + y' \cos\theta$$

d'où, pour les dérivées partielles :

$$\partial_{x'} = \cos\theta\, \partial_x + \sin\theta\, \partial_y \quad \text{et} \quad \partial_{y'} = -\sin\theta\, \partial_x + \cos\theta\, \partial_y .$$

Ces relations sont identiques à celles qui font passer de (x, y) à (x', y'). Plus généralement si $\underline{\mathrm{d}x} = \mathbf{M}\, \underline{\mathrm{d}x'}$ ($\mathbf{M}$ matrice $n \times p$), alors $\partial_{\underline{x}'} = \mathbf{M}^t\, \partial_{\underline{x}}$ ($\mathbf{M}^t$ transposée de $\mathbf{M}$). Donc $\partial_{\underline{x}}$ a la même loi de transformation que $\underline{\mathrm{d}x}$ si $\mathbf{M}^t = \mathbf{M}^{-1}$ (matrice orthogonale $n \times n$, par exemple une rotation pour $n = 3$).

EXEMPLE 3. Dans une **transformation de Lorentz**, $ct = ct' \cosh\varphi + x' \sinh\varphi$ et $x = ct' \sinh\varphi + x' \cosh\varphi$, on vérifie que $(\frac{1}{c}\partial_t, \partial_x)$ se transforme comme $(ct, -x)$; plus généralement $(\frac{1}{c}\partial_t, -\overrightarrow{\nabla})$ se transforme comme le **quadrivecteur** $(ct, \vec{r})$. Dans une **transformation de Galilée** : $t = t'$ et $\vec{r} = \vec{r'} + \overrightarrow{V}t'$, d'où : $\partial_{t'} = \partial_t + \overrightarrow{V} \cdot \overrightarrow{\nabla} \neq \partial_t$ (donc stationnarité $\partial_t = 0$ dépendante du référentiel), $\overrightarrow{\nabla}' = \overrightarrow{\nabla}$, et $\mathrm{d}_{t'} \overset{\text{def}}{=} \partial_{t'} + \vec{v'} \cdot \overrightarrow{\nabla}' = \partial_t + \overrightarrow{V} \cdot \overrightarrow{\nabla} + (\vec{v} - \overrightarrow{V}) \cdot \overrightarrow{\nabla} = \mathrm{d}_t$ (dérivée particulaire invariante).

7.2 CALCUL INTÉGRAL

7.2.1 Intégration à n dimensions ; jacobien ; cas des très grandes dimensions

▷ Intégration et domaines d'intégration

Rappelons que le volume dans $\mathbb{R}^n$ d'un "**parallélépipède de base**" $\mathcal{P}$ (produit des intervalles $[a_i, b_i]$) est $\mu(\mathcal{P}) = (b_1 - a_1)(b_2 - a_2)\cdots(b_n - a_n)$. Dans l'approche de Riemann l'intégrale $\iint\cdots\int_{\mathcal{D}} f(x_1, x_2 \cdots x_n)\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_n$ est définie, à partir du pavage de $\mathbb{R}^n$ avec des parallélépipèdes de base $\mathcal{P}_\alpha$, comme la limite lorsque les $\mu(\mathcal{P}_\alpha)$ tendent vers zéro de la somme $\sum_\alpha f(\underline{x}_\alpha)\,\mu(\mathcal{P}_\alpha)$ étendue aux $\mathcal{P}_\alpha$ qui intersectent $\mathcal{D}$ ($\underline{x}_\alpha$ point de $\mathcal{P}_\alpha$). En pratique le calcul d'une telle intégrale se ramène à une suite de calculs d'intégrales à une dimension (sur x_i à $x_{k(\neq i)}$ fixés puis sur x_j à $x_{k(\neq i \text{ et } j)}$ fixés, *etc.*). Le résultat ne dépend pas, lorsque l'intégrale a un sens, de l'ordre dans lequel sont effectuées les intégrations successives ; par contre les **domaines** des intégrations partielles eux en dépendent. Par exemple (figure 9) :

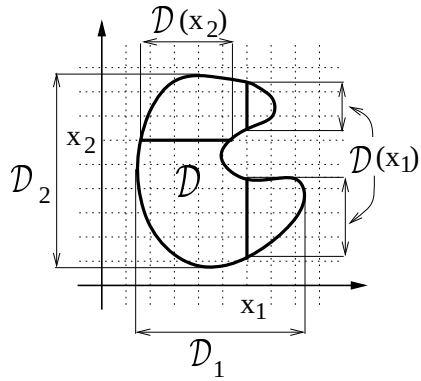


FIGURE 9

$$\iint_{\mathcal{D}} f(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = \int_{\mathcal{D}_1} \Big[\int_{\mathcal{D}(x_1)} f(x_1, x_2)\,\mathrm{d}x_2\Big]\,\mathrm{d}x_1 = \int_{\mathcal{D}_2} \Big[\int_{\mathcal{D}(x_2)} f(x_1, x_2)\,\mathrm{d}x_1\Big]\,\mathrm{d}x_2 \ ;$$

même si $f(x_1, x_2) = f_1(x_1) f_2(x_2)$, le résultat n'est pas $\displaystyle\Big[\int_{\mathcal{D}_1} f_1(x_1)\,\mathrm{d}x_1\Big] \times \Big[\int_{\mathcal{D}_2} f_2(x_2)\,\mathrm{d}x_2\Big]$ (sauf si $\mathcal{D} = \mathcal{D}_1 \times \mathcal{D}_2$). Heureusement le domaine $\mathcal{D}$ et la fonction f possèdent souvent

des propriétés de symétrie et, par un choix judicieux de nouvelles variables d'intégration, on peut se ramener à un domaine simple (parallélépipède de base) ; reste à savoir récrire une intégrale lorsqu'on fait un changement de variables.

▷ Changements de variables $\underline{x}(\underline{x}')$

Le pavage avec des parallélépipèdes infinitésimaux de base pour les coordonnées $\underline{x}'$ correspond, pour les coordonnées $\underline{x}$, à un pavage dont les parallélépipèdes ne sont pas de base (figures 10 et 11).

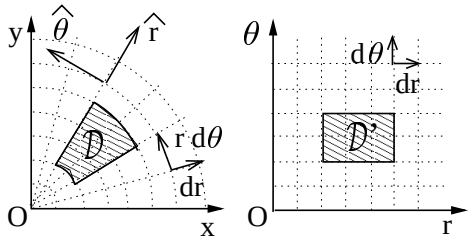


FIGURE 10

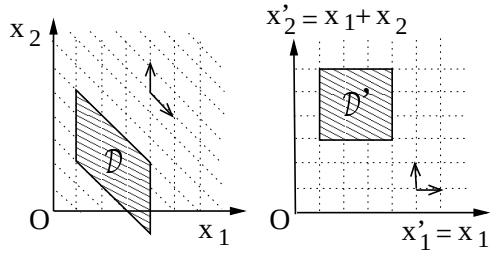


FIGURE 11

A chaque variation $\mathrm{d}x'_j$ (j fixé) correspondent les variations $\dfrac{\partial x_i}{\partial x'_j}\,\mathrm{d}x'_j$ des x_i et donc, dans l'espace $\underline{x}$, un déplacement $\delta_j\underline{x} = \displaystyle\sum_i \left(\frac{\partial x_i}{\partial x'_j}\,\mathrm{d}x'_j\right)\hat{e}_i$ qui n'est pas parallèle à un vecteur de base $\hat{e}_k$ de cet espace. Le volume (algébrique) du parallélépipède construit sur les vecteurs $\delta_j\underline{x}$ vaut (*cf.* section 4.1.3) $J(\underline{x},\underline{x}')\,\mathrm{d}x'_1\,\mathrm{d}x'_2\cdots\mathrm{d}x'_n$, où $J(\underline{x},\underline{x}')$ est le déterminant (**jacobien**) de la matrice (de Jacobi) d'éléments $\dfrac{\partial x_i}{\partial x'_j}$. Avec ce nouveau pavage :

$$\iint\cdots\int_{\mathcal{D}} f(\underline{x})\,\mathrm{d}x_1\,\mathrm{d}x_2\cdots\mathrm{d}x_n = \iint\cdots\int_{\mathcal{D}'} f\big(\underline{x}(\underline{x}')\big)\,|J(\underline{x},\underline{x}')|\,\mathrm{d}x'_1\,\mathrm{d}x'_2\cdots\mathrm{d}x'_n\ ,$$

où $\mathcal{D}'$ est le domaine (en $\underline{x}'$) associé à $\mathcal{D}$ (en $\underline{x}$).

REMARQUE : si $J(\underline{x}',\underline{x})$ est plus facile à calculer que $J(\underline{x},\underline{x}')$, on utilise $J(\underline{x},\underline{x}') = [J(\underline{x}',\underline{x})]^{-1}$.

EXEMPLE 1. $x_1 = x'_1$ et $x_2 = ax'_2 - x'_1$; alors $\delta_1\underline{x} = \mathrm{d}x'_1(\hat{e}_1 - \hat{e}_2)$ et $\delta_2\underline{x} = a\,\mathrm{d}x'_2\,\hat{e}_2$; $J = \det\begin{pmatrix} 1 & 0 \\ -1 & a \end{pmatrix} = a$.

EXEMPLE 2. Si $\underline{x} = \vec{r}$ et si les $\delta_j\vec{r}$ forment un trièdre orthonormal direct, le volume infinitésimal est le produit des $|\delta_j\vec{r}|$, égal par exemple à $(\mathrm{d}r)\,(r\,\mathrm{d}\theta)\,(r\,\sin\theta\,\mathrm{d}\varphi)$ en coordonnées sphériques.

Exemple 3. Le potentiel rayonné par une distribution de charge quasi-ponctuelle (de charge totale q) suivant une trajectoire $\vec{s}(t)$ (parcourue à la vitesse $\vec{v}(t) = \dot{\vec{s}}(t)$) s'écrit (*cf.* section 7.5.2) :

$$V(\vec{r}_0,t_0)=\tfrac{1}{4\pi\epsilon_0}\iiint\frac{\rho\left(\vec{r}-\vec{s}\left(t_0-\frac{|\vec{r}-\vec{r}_0|}{c}\right)\right)}{|\vec{r}-\vec{r}_0|}\,\mathrm{d}^3r\ .$$

Bien que le domaine d'intégration soit concentré sur la charge et se limite donc au seul point $\vec{s}^*=\vec{s}(t^*)$, position de la charge à l'instant retardé défini par $t^*=t_0-\frac{|\vec{s}(t^*)-\vec{r}_0|}{c}$, le résultat n'est pas $V(\vec{r}_0,t_0)=\frac{q}{4\pi\epsilon_0 R^*}$ (avec $\overrightarrow{R}^*=\vec{r}_0-\vec{s}^*=R^*\hat{u}^*$) mais :

$$V(\vec{r}_0,t_0)=\tfrac{q}{4\pi\epsilon_0 R^*}\left(1-\tfrac{\vec{v}(t^*)\cdot\hat{u}^*}{c}\right)^{-1}\quad\textbf{(potentiel de Lienart-Wiechert)}\ .$$

La raison est que le jacobien qui correspond au changement de variable $\vec{r}\,'=\vec{r}-\vec{s}\big(t_0-\frac{|\vec{r}-\vec{r}_0|}{c}\big)$ est (petit calcul) $J(\vec{r}\,',\vec{r})=\det\big|\delta_{ij}-\frac{v_i u_j}{c}\big|=1-\frac{\vec{v}\cdot\hat{u}}{c}$. On retrouve ce résultat plus rapidement en écrivant :

$$V(\vec{r}_0,t_0)=\tfrac{1}{4\pi\epsilon_0}\iiint\mathrm{d}^3r\iint\frac{\rho(\vec{r}-\vec{s}(t))}{|\vec{r}-\vec{r}_0|}\,\delta\left(t-t_0+\tfrac{|\vec{r}-\vec{r}_0|}{c}\right)\mathrm{d}^3r\,\mathrm{d}t\ .$$

L'intégration sur $\vec{r}$ donne $V(\vec{r}_0,t_0)=\frac{q}{4\pi\epsilon_0}\int|\vec{s}(t)-\vec{r}_0|^{-1}\,\delta\big(t-t_0+\frac{|\vec{s}(t)-\vec{r}_0|}{c}\big)\,\mathrm{d}t$ et celle sur t, avec le changement de variable $t'=t-t_0+\frac{|\vec{s}(t)-\vec{r}_0|}{c}$, conduit au résultat (car $\frac{\mathrm{d}t'}{\mathrm{d}t}=1-\frac{\vec{v}\cdot\hat{u}}{c}$).

▷ Volumes à très grand nombre de dimensions

L'intuition des petites dimensions est trompeuse lorsque n devient grand. Exemples : le volume d'un domaine $\mathcal{D}_{n,\alpha}(R)$ défini par $\sum_{i=1}^n|x_i|^\alpha<R^\alpha$ (sphère "pleine" si $\alpha=2$) devient concentré sur sa surface et la loi de répartition d'une coordonnée devient "piquée" en 0 ; le "rayon" typique d'un cube de côté a (volume a^n) est proportionnel à $a\sqrt{n}$.

DEMONSTRATION. Le volume de $\mathcal{D}_{n,\alpha}(R)$ s'écrit dimensionnellement $V_{n,\alpha}(R)=R^n\,V_{n,\alpha}$, d'où pour n grand (*cf.* section 1.1.4 : $n=10^{14}$ et $\frac{\delta R}{R}=10^{-6}$) $\frac{V_{n,\alpha}(R-\delta R)}{V_{n,\alpha}(R)}=\left(1-\frac{\delta R}{R}\right)^n\simeq\exp\left(-n\frac{\delta R}{R}\right)\simeq 0$ (au lieu de $1-n\frac{\delta R}{R}\simeq 1$ pour n petit). Si on fixe par exemple x_1, le nouveau domaine défini par $\sum_{i=2}^n|x_i|^\alpha<R^\alpha-|x_1|^\alpha$ a un volume proportionnel à $(R^\alpha-|x_1|^\alpha)^{\frac{n-1}{\alpha}}$, qui se comporte pour n grand comme $R^{n-1}\exp\left(-\frac{n}{\alpha}\frac{|x_1|^\alpha}{R^\alpha}\right)$. Si toutes les positions à l'intérieur de $\mathcal{D}_{n,\alpha}(R)$ sont supposées équiprobables, cette quantité est proportionnelle (*cf.* section 10.1.2) à la densité de probabilité $P_{n,\alpha}(x_1)$ de x_1 (figure 12). Pour le cube centré en O : $R^2=<\sum_{i=1}^n x_i^2>=n\frac{a^2}{12}$

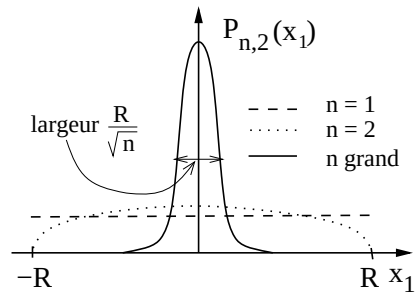


FIGURE 12

REMARQUE. Le calcul de $V_{n,\alpha}$ (volume pour $R=1$) peut se faire en écrivant

$$1=\textstyle\prod_{i=1}^n\left(\int_{-\infty}^{\infty}e^{-A|x_i|^\alpha}\,\mathrm{d}x_i\right)=\int_0^\infty e^{-Ar^\alpha}\,n\,V_{n,\alpha}\,r^{n-1}\,\mathrm{d}r\ .$$

$A = 2$ si $\alpha = 1$ et $A = \pi$ si $\alpha = 2$. A l'aide de la méthode de dérivation par rapport à un paramètre décrite à la section 5.1.3., on obtient exactement $V_{n,1} = \frac{2^n}{n!}$, $V_{2p,2} = \frac{\pi^p}{p!}$ et $V_{2p+1,2} = \pi^p \, \frac{2^{p+1}}{(2p+1)!!}$. Pour α quelconque et n grand, le changement de variable $r = n^{\frac{1}{\alpha}} u$ donne le comportement dominant

$$V_{n,\alpha}(R) \sim R^n \, n^{-\frac{n}{\alpha}}$$

(au lieu de R^n pour un cube de côté R). Le calcul de l'intégrale restante, dans l'approximation de la méthode du col, ne donne que des facteurs "sous dominants" du type $\exp(\mathrm{C}^{\mathrm{ste}} n)$.

EXEMPLE (*cf.* section 1.1.4). En physique de tels volumes interviennent dans l'espace de phase en physique statistique ; par exemple la région $U - \delta U \leq \sum_{i=1}^{N} \frac{1}{2}\big(\frac{\vec{p}_i^{\,2}}{m} + K\vec{r}_i^{\,2}\big) \leq U$ à $6N$ dimensions pour N oscillateurs ($\alpha = 2$) a un volume d'ordre $\big(\frac{U}{N}\big)^{3N}$, et la densité de probabilité $P_{6N,2}(x_i$ ou $p_{i,x})$ n'est autre que la **loi de Boltzmann** (réflexion laissée au lecteur).

7.2.2 Formes différentielles ; formules de Stokes et Ostrogradski

▷ Forme différentielle

C'est une expression $\sum_i A_i(\underline{x})\,\mathrm{d}x_i$ qui sert à définir des intégrales curvilignes le long de chemins γ paramétrés :

$$\int_\gamma \sum_i A_i(\underline{x})\,\mathrm{d}x_i \overset{\text{def}}{=} \int_{\lambda_1}^{\lambda_2} \sum_i A_i\big(\underline{x}(\lambda)\big)\,\frac{\mathrm{d}x_i}{\mathrm{d}\lambda}\,\mathrm{d}\lambda \ .$$

Si le chemin est fermé ($\underline{x}(\lambda_1) = \underline{x}(\lambda_2)$) l'intégrale curviligne est notée $\oint_\gamma$. La différentielle d'une fonction f est un cas particulier pour lequel :

$$A_i(\underline{x}) = \frac{\partial f}{\partial x_i}(\underline{x}) \quad (1) \quad \text{d'où} \quad \frac{\partial A_i}{\partial x_j} = \frac{\partial A_j}{\partial x_i} \quad (2) \quad \text{et} \quad \oint_\gamma \sum_i A_i(\underline{x})\,\mathrm{d}x_i = 0 \quad (3) \ .$$

Inversement si (3) est vérifiée pour tout chemin γ fermé dans un domaine $\mathcal{D}$, la forme $\sum_i A_i(\underline{x})\,\mathrm{d}x_i$ est bien la différentielle d'une fonction (unique à une constante additive près). Elle est définie par $f(\underline{x}) = \int_{\underline{x}_0}^{\underline{x}} \sum_i A_i(\underline{x}')\,\mathrm{d}x'_i + f(\underline{x}_0)$, car alors l'intégrale de $\underline{x}_0$ à $\underline{x}$ ne dépend pas du chemin suivi (figure 13).

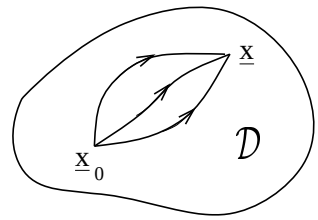


FIGURE 13

On va voir, sur les cas particuliers $n = 2$ et $n = 3$, que (2) implique aussi (1) sous la réserve importante que, dans $\mathcal{D}$, tout chemin fermé peut être réduit continûment à un point.

▷ Formule de Stokes (ou Green) dans le plan

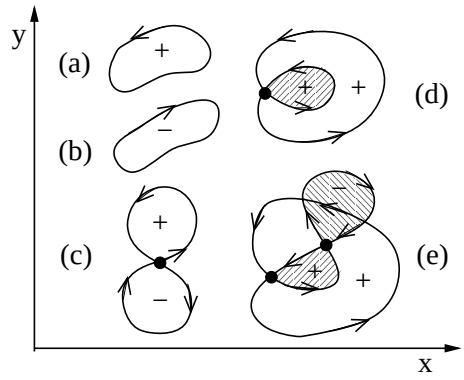


FIGURE 14

Dans le cas simple d'un chemin fermé ne se coupant pas, limitant une surface S et parcouru soit dans le sens direct (figure 14a) soit dans le sens indirect (figure 14b), on a

$$\oint_\gamma (A_x\,\mathrm{d}x + A_y\,\mathrm{d}y) = \iint_S (\partial_x A_y - \partial_y A_x)\,\mathrm{d}S \quad \text{avec} \;\; \mathrm{d}S = \pm \mathrm{d}x\,\mathrm{d}y$$

(signe + pour le sens direct et − pour le sens indirect).

DÉMONSTRATION : pour un rectangle "direct" (figure 15a) la circulation vaut exactement

$$\int_{x_1}^{x_2} \big(A_x(x,y_1) - A_x(x,y_2)\big)\,\mathrm{d}x + \int_{y_1}^{y_2} \big(A_y(x_2,y) - A_x(x_1,y)\big)\,\mathrm{d}y = \int_{x_1}^{x_2}\int_{y_1}^{y_2} \big(-\partial_y A_x + \partial_x A_y\big)\,\mathrm{d}x\,\mathrm{d}y\ .$$

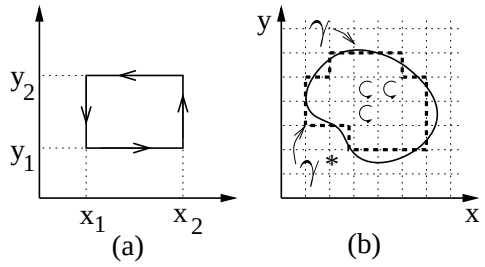


FIGURE 15

Pour étendre ce résultat au chemin "direct" γ de la figure 15b, il suffit de paver le plan avec des rectangles arbitrairement petits permettant d'approcher γ par un chemin en escalier γ^*. La somme des circulations sur les rectangles est égale à celle sur γ^*, qui dans la limite tend vers celle sur γ. La généralisation de ce résultat à un circuit fermé

γ quelconque (figures 14c,d,e) s'obtient en remarquant qu'il est la réunion de circuits fermés simples, et en ajoutant leurs contributions.

REMARQUE (figure 16). La démonstration demande que $\overrightarrow{A}$ et ses dérivées soient définies à l'intérieur de γ. Un contre exemple connu est celui du champ $\overrightarrow{B}$ créé par un fil infini (selon z) parcouru par un courant I, pour lequel $\partial_x B_y - \partial_y B_x = 0$ mais $\oint_\gamma \overrightarrow{B} \cdot \overrightarrow{\mathrm{d}l} = \mu_0 I$ si γ entoure le fil ; $\overrightarrow{B}$ n'étant pas défini sur le fil, γ ne peut être réduit à un point. Un autre est l'écoulement irrotationnel d'un fluide autour d'une aile d'avion (*cf.* section 8.3.3).

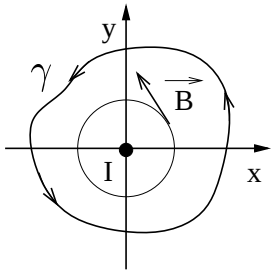


FIGURE 16

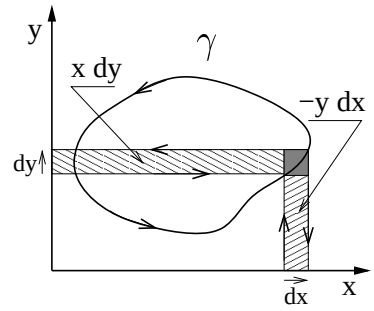


FIGURE 17

EXEMPLE (figure 17). L'**aire algébrique** associée à une courbe fermée orientée γ (*cf.* section 2.2.1) est :

$$S = \oint_\gamma x\,\mathrm{d}y = \oint_\gamma -y\,\mathrm{d}x = \frac{1}{2}\oint_\gamma (x\,\mathrm{d}y - y\,\mathrm{d}x)\ .$$

Si $x = V$ (volume) et $y = P$ (pression), $\oint_\gamma -P\,\mathrm{d}V$ est le travail reçu au cours d'un cycle (positif si le cycle est direct).

▷ Formule de Stokes à trois dimensions (figure 18)

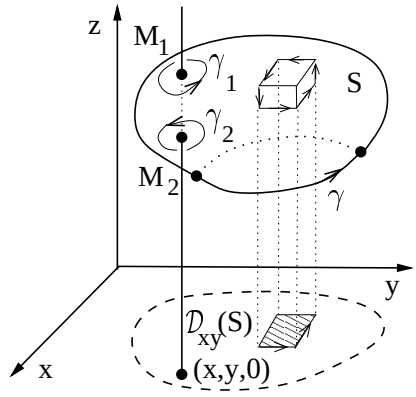


FIGURE 18

Soit γ une courbe fermée orientée et S une surface ayant γ pour frontière. On suppose de plus que S est une **surface orientable**, c.a.d. que pour tout point $\vec{r}$ de S tout circuit élémentaire $\gamma_{\vec{r}}$ entourant ce point possède une orientation qui peut être définie de manière

unique "par continuité" à partir de celle de γ ; (ceci n'est pas toujours possible ; par exemple, en un tour sur la bande de Möbius, γ est changé en $-\gamma$; figure 19).

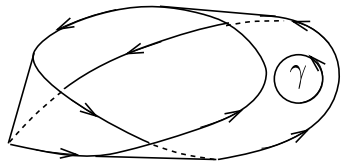


FIGURE 19

On a alors :

$$\oint_\gamma (A_x\,\mathrm{d}x + A_y\,\mathrm{d}y + A_z\,\mathrm{d}z) = \iint_{\mathcal{D}_{yz}(S)} (\partial_y A_z - \partial_z A_y)\,\mathrm{d}S_x + \iint_{\mathcal{D}_{zx}(S)} (\partial_z A_x - \partial_x A_z)\,\mathrm{d}S_y + \iint_{\mathcal{D}_{xy}(S)} (\partial_x A_y - \partial_y A_x)\,\mathrm{d}S_z \ .$$

Les domaines d'intégration des intégrales doubles sont les projections de S sur les **plans orientés** (y,z), (z,x) et (x,y), et le choix de signe $\mathrm{d}S_x = \pm\mathrm{d}y\,\mathrm{d}z$, $\mathrm{d}S_y = \pm\mathrm{d}z\,\mathrm{d}x$ et $\mathrm{d}S_z = \pm\mathrm{d}x\,\mathrm{d}y$ en un point $\vec{r}$ dépend du caractère direct ou indirect de la projection de $\gamma_{\vec{r}}$ sur ces plans.

DÉMONSTRATION. Le résultat est évident si γ suit les côtés d'un parallélépipède de base (figure 20).

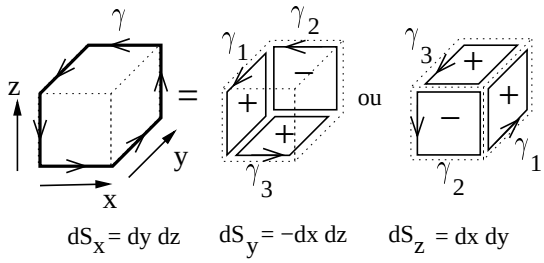


FIGURE 20

En effet on a $\oint_\gamma = \oint_{\gamma_1} + \oint_{\gamma_2} + \oint_{\gamma_3}$ où les γ_i sont les projections de γ sur les plans (y,z), (z,x) et (x,y) et à chaque γ_i on peut appliquer la formule de Stokes à deux dimensions (puisque la troisième variable est fixée). On étend ce résultat par un pavage de l'espace permettant d'approcher la surface S (figure 18) par une surface "en escaliers" S^* constituée par un ensemble de rectangles arbitrairement petits parallèles aux plans de base (l'orientabilité assurant que les circulations sur ces rectangles se compensent, sauf sur γ). Pour les rectangles parallèles au plan de base (x,y), par exemple, la somme des circulations donne l'intégrale sur S_z. Il y a en général plusieurs points M de S de coordonnées (x,y) ; pour chacun on calcule $\partial_x A_y - \partial_y A_x$ et $\mathrm{d}S_z = \pm\mathrm{d}x\,\mathrm{d}y$ suivant l'orientation de la projection de $\gamma_{\vec{r}}$ sur le plan (x,y).

REMARQUE : aucune structure métrique n'a été utilisée dans la démonstration. Cependant lorsqu'elle existe (cas de l'**espace usuel**), le choix $\mathrm{d}S_z = \pm\mathrm{d}x\,\mathrm{d}y$ correspond simplement au signe du produit scalaire $\overrightarrow{\mathrm{d}S}\cdot\hat{z}$ avec $\overrightarrow{\mathrm{d}S} = \mathrm{d}S\,\hat{n}$, le sens de $\hat{n}$ étant déduit de l'orientation de $\gamma_{\vec{r}}$ par la **règle du tire-bouchon**. On a alors $\oint_\gamma \overrightarrow{A}\cdot\overrightarrow{\mathrm{d}r} = \iint_S \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}\cdot\overrightarrow{\mathrm{d}S}$.

▷ **Formule d'Ostrogradski** (espace à trois dimensions)

Pour un volume V limité par une surface S fermée, **orientée vers l'extérieur du volume**, on a :

$$\iint_S (A_x\,\mathrm{d}S_x + A_y\,\mathrm{d}S_y + A_z\,\mathrm{d}S_z) = \iiint_V \left[\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right] \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ .$$

DÉMONSTRATION. Pour le parallélépipède de la figure 21, et les orientations choisies pour ses faces, on écrit : $\iiint (\partial_z A_z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{x_1}^{x_2}\int_{y_1}^{y_2} [A_z(x,y,z_2) - A_z(x,y,z_1)]\,\mathrm{d}x\,\mathrm{d}y = \iint_S A_z(x,y,z)\,\mathrm{d}S_z$; de même $\iiint (\partial_{x\,(\text{ou}\,y)} A_{x\,(\text{ou}\,y)}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iint_S A_{x\,(\text{ou}\,y)}(x,y,z)\,\mathrm{d}S_{x\,(\text{ou}\,y)}$. Le résultat général s'obtient par un pavage de l'espace par des parallélépipèdes arbitrairement petits.

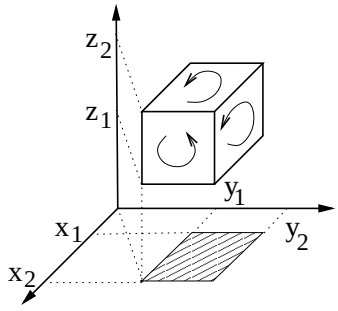


FIGURE 21

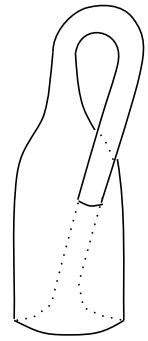


FIGURE 22

REMARQUE : il existe des surfaces fermées (non orientables) qui ne limitent pas de volume (bouteille de Klein ; figure 22).

▷ **Expressions de $\overrightarrow{\mathrm{grad}} f$, $\mathrm{div}\,\vec{A}$, $\overrightarrow{\mathrm{rot}}\,\vec{A}$ et Δf en coordonnées curvilignes**

Ces formules, d'une utilité limitée puisque le cas de champs symétriques a déjà été traité à la section 7.2.1, sont plus faciles à obtenir en considérant plus généralement un système de coordonnées curvilignes $(x_1, x_2, x_3) \neq (x, y, z)$ tel que les vecteurs $\hat{e}_1$, $\hat{e}_2$ et $\hat{e}_3$ définis par $\frac{\partial \vec{r}}{\partial x_i} = e_i\,\hat{e}_i$ forment une base orthonormée directe. Dans cette base, on a (en tout point $\vec{r}$) :

$$\vec{A} = A_1\hat{e}_1 + A_2\hat{e}_2 + A_3\hat{e}_3 \quad , \quad \vec{\mathrm{d}r} = e_1\,\mathrm{d}x_1\,\hat{e}_1 + e_2\,\mathrm{d}x_2\,\hat{e}_2 + e_3\,\mathrm{d}x_3\,\hat{e}_3\ ,$$
$$\vec{\mathrm{d}S} = e_2e_3\,\mathrm{d}x_2\,\mathrm{d}x_3\,\hat{e}_1 + e_3e_1\,\mathrm{d}x_3\,\mathrm{d}x_1\,\hat{e}_2 + e_1e_2\,\mathrm{d}x_1\,\mathrm{d}x_2\,\hat{e}_3 \qquad \mathrm{d}^3r = e_1e_2e_3\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3\ .$$

Alors des formules

$$\mathrm{d}f = \textstyle\sum_i \partial_i f\,\mathrm{d}x_i = \overrightarrow{\mathrm{grad}} f \cdot \vec{\mathrm{d}r}\ ,$$
$$\oint \vec{A}\cdot\vec{\mathrm{d}r} = \oint (A_1e_1\,\mathrm{d}x_1 + \cdots) = \iint ((\partial_2(e_3A_3) - \partial_3(e_2A_2))\,\mathrm{d}x_2\,\mathrm{d}x_3 + \cdots) = \iint \overrightarrow{\mathrm{rot}}\,\vec{A}\cdot\vec{\mathrm{d}S}\ (\textit{cf.}\ \text{Stokes})\ ,$$
$$\oiint \vec{A}\cdot\vec{\mathrm{d}S} = \oiint (A_1e_2e_3\,\mathrm{d}x_2\,\mathrm{d}x_3 + \cdots) = \iiint (\partial_1(e_2e_3A_1) + \cdots)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 = \iiint \mathrm{div}\,\vec{A}\,\mathrm{d}^3r$$

(*cf.* Ostrogradski), et enfin de $\Delta f = \mathrm{div}\,(\overrightarrow{\mathrm{grad}} f)$, on déduit respectivement les expressions

$$\overrightarrow{\mathrm{grad}} f = \tfrac{1}{e_1}\tfrac{\partial f}{\partial x_1}\,\hat{e}_1 + \tfrac{1}{e_2}\tfrac{\partial f}{\partial x_2}\,\hat{e}_2 + \tfrac{1}{e_3}\tfrac{\partial f}{\partial x_3}\,\hat{e}_3\ ,$$
$$\overrightarrow{\mathrm{rot}}\,\vec{A} = \tfrac{1}{e_2e_3}\Big(\tfrac{\partial}{\partial x_2}(e_3A_3) - \tfrac{\partial}{\partial x_3}(e_2A_2)\Big)\,\hat{e}_1 + \text{ permutation circulaire}\ ,$$
$$\mathrm{div}\,\vec{A} = \tfrac{1}{e_1e_2e_3}\Big(\tfrac{\partial}{\partial x_1}(e_2e_3A_1) + \tfrac{\partial}{\partial x_2}(e_3e_1A_2) + \tfrac{\partial}{\partial x_3}(e_1e_2A_3)\Big)\ ,$$

$$\Delta f = \frac{1}{e_1 e_2 e_3}\left(\frac{\partial}{\partial x_1}\left(\frac{e_2 e_3}{e_1}\,\frac{\partial f}{\partial x_1}\right) + \frac{\partial}{\partial x_2}\left(\frac{e_3 e_1}{e_2}\,\frac{\partial f}{\partial x_2}\right) + \frac{\partial}{\partial x_3}\left(\frac{e_1 e_2}{e_3}\,\frac{\partial f}{\partial x_3}\right)\right) .$$

Pour les **coordonnées sphériques** (r,θ,φ) (resp. **cylindriques** (r,θ,z)), on a $(e_1,e_2,e_3)\equiv(1,r,r\sin\theta)$ (resp. $(1,r,1)$).

7.2.3 Analyse vectorielle ; frontières et champs dépendant du temps ; loi de Lenz ; hydrodynamique

L'analyse vectorielle repose sur les identités (1-3), ci-dessous et les propriétés de $\overrightarrow{\nabla}$.

▷ Identités de base ; vocabulaire et notations associés

$$f(\vec r_2)-f(\vec r_1)=\int_1^2 \overrightarrow{\mathrm{grad}} f\cdot \mathrm{d}\vec r \quad (1) \quad , \quad \oint_\gamma \overrightarrow{A}\cdot \mathrm{d}\vec r = \iint_S \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}\cdot \mathrm{d}\overrightarrow{S} \quad (2) \quad \text{et}$$

$$\oiint_S \overrightarrow{A}\cdot \mathrm{d}\overrightarrow{S} = \iiint_V \mathrm{div}\,\overrightarrow{A}\,\mathrm{d}^3 r \quad (3) .$$

(1) a été discutée à la section 7.1.2. (2) (Stokes) relie la **circulation** $\mathcal{C}_\gamma(\overrightarrow{A})$ du champ $\overrightarrow{A}$ le long du contour γ au **flux** $\varphi_S(\overrightarrow{\mathrm{rot}}\,\overrightarrow{A})$ **à travers une surface ouverte** S limitée et orientée par γ ; si γ est infinitésimal $\mathcal{C}_\gamma(\overrightarrow{A}) = \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}\cdot\mathrm{d}\overrightarrow{S}$. (3) (Ostrogradski) relie le **flux** $\Phi_S(\overrightarrow{A})$ de $\overrightarrow{A}$ **à travers une surface fermée** S (orientée vers l'extérieur) à l'intégrale, dans le volume V qu'elle entoure, de $\mathrm{div}\,\overrightarrow{A}$; si V est infinitésimal $\Phi_S(\overrightarrow{A}) = \mathrm{div}\,\overrightarrow{A}\,\mathrm{d}^3 r$. Avec la convention de sommation d'Einstein (*cf.* section 4.1.3), (2) et (3) s'écrivent

$$\oint_\gamma A_i\,\mathrm{d}x_i = \iint_S \epsilon_{ijk}\,\partial_j A_k\,\mathrm{d}S_i \quad (2') \quad \text{et} \quad \oiint_S A_i\,\mathrm{d}S_i = \iiint_V (\partial_i A_i)\,\mathrm{d}^3 r \quad (3') ,$$

où $A_i \equiv A_x, A_y, A_z$ peuvent être trois fonctions arbitraires.

Conditions de nullité. 1) Si $\overrightarrow{A} = \overrightarrow{\mathrm{grad}} f$ (resp. $\overrightarrow{A} = \overrightarrow{\mathrm{rot}}\,\overrightarrow{V}$), les intégrales (2) (resp. (3)) sont nulles. La réciproque est vraie si le domaine dans lequel les intégrales sont nulles peut être continûment réduit à un point. (Remarque : on peut avoir à la fois $\overrightarrow{A} = \overrightarrow{\nabla} f$ et $\overrightarrow{A} = \overrightarrow{\nabla}\wedge\overrightarrow{V}$, par exemple $f(\vec r) = \overrightarrow{A}\cdot\vec r$ et $\overrightarrow{V} = \dfrac{1}{2}\,\overrightarrow{A}\wedge\vec r$). 2) L'intégrale (2) (resp. (3)) tend vers 0 si γ (resp. S) s'en va à l'infini et si $|\overrightarrow{A}(\vec r)|$ décroit plus vite que r^{-1} (resp. r^{-2}) lorsque $r\to\infty$.

▷ Autres identités

Celles-ci étant très nombreuses on se limite à deux exemples.

EXEMPLE 1. Dans un milieu continu la résultante des **forces de contraintes** à la surface S d'un volume V et leur puissance s'écrivent comme des intégrales sur V :

$$\oiint_S [\boldsymbol{\sigma}]\,\mathrm{d}\overrightarrow{S} = \iiint_V \vec f_c\,\mathrm{d}^3 r \quad \text{avec } f_{c,x} = \partial_x\sigma_{xx} + \partial_y\sigma_{xy} + \partial_z\sigma_{xz} \text{ (idem pour } f_{c,y} \text{ et } f_{c,z})$$

$$\text{et} \quad \oiint_S \vec v\cdot[\boldsymbol{\sigma}]\mathrm{d}\overrightarrow{S} = \iiint_V \mathcal{P}\,\mathrm{d}^3 r \quad \text{avec } \mathcal{P} = \mathrm{div}([\boldsymbol{\sigma}]\vec v) = \vec f_c\cdot\vec v + \mathrm{tr}\left([\boldsymbol{\sigma}][\dot{\boldsymbol{\epsilon}}]\right) .$$

($[\boldsymbol{\sigma}] = -[\mathbf{P}]$ est la matrice des contraintes ; *cf.* section 4.3.1). Ceci donne pour des forces de pression :

$$\oiint_S -P\,\overrightarrow{\mathrm{d}S} = \iiint_V -\overrightarrow{\mathrm{grad}}\,P\,\mathrm{d}^3r \quad ; \quad \mathcal{P} = -\overrightarrow{\mathrm{grad}}\,P\cdot\vec{v} - P\,\mathrm{div}\vec{v}\,.$$

Démonstration : $\oiint_S \sigma_{ij}\,\mathrm{d}S_j = \iiint_V \partial_j\,\sigma_{ij}\,\mathrm{d}^3r$ et $\oiint_S v_i\sigma_{ij}\,\mathrm{d}S_j = \iiint_V \partial_j\,(v_i\sigma_{ij})\,\mathrm{d}^3r = \iiint_V v_i\,\partial_j\,\sigma_{ij}\,\mathrm{d}^3r + \iiint_V \sigma_{ij}\,\frac{1}{2}\,(\partial_i v_j + \partial_j\,v_i)\,\mathrm{d}^3r$ car σ_{ij} est symétrique. (Si $[\boldsymbol{\sigma}]$ avait une partie antisymétrique c_{ij}, il faudrait rajouter l'intégrale de $c_{ji}\,\frac{1}{2}\,(\partial_i v_j - \partial_j\,v_i)$ interprétable comme la puissance $\overrightarrow{C}\cdot\vec{\omega}$ de couples volumiques ; l'hypothèse $\sigma_{ij} = \sigma_{ji}$ équivaut donc à l'absence de tels couples dans la modélisation habituelle des milieux continus.)

EXEMPLE 2. Dans le calcul du potentiel scalaire (ou vectoriel) dû à une répartition volumique de **dipoles électriques** (ou **magnétiques**), on utilise les relations (avec $f(\vec{r}) \propto |\vec{r}-\vec{r}_0|^{-1}$) :

$\iiint_V (\overrightarrow{P}\cdot\overrightarrow{\nabla}f$ ou $\overrightarrow{M}\wedge\overrightarrow{\nabla}f)\,\mathrm{d}^3r = \iiint_V f\,(-\overrightarrow{\nabla}\cdot\overrightarrow{P}$ ou $\overrightarrow{\nabla}\wedge\overrightarrow{M})\,\mathrm{d}^3r + \oiint_S f\,(\overrightarrow{P}\cdot\overrightarrow{\mathrm{d}S}$ ou $\overrightarrow{M}\wedge\overrightarrow{\mathrm{d}S})$.

Démonstration : on considère les intégrales de surface. Dans le cas électrique $\Phi_S(f\overrightarrow{P})$ est l'intégrale sur V de $\mathrm{div}(f\overrightarrow{P}) = \overrightarrow{P}\cdot\overrightarrow{\nabla}f + f\,\mathrm{div}\overrightarrow{P}$. Dans le cas magnétique, quel que soit $\vec{u}$ constant, $\vec{u}\oiint_S f\overrightarrow{M}\wedge\overrightarrow{\mathrm{d}S} = \Phi_S(f\vec{u}\wedge\overrightarrow{M})$ ($\vec{u}$ "passant sous le signe somme") est l'intégrale sur V de :

$$\mathrm{div}(f(\vec{u}\wedge\overrightarrow{M})) = f(\overrightarrow{M}\cdot\overrightarrow{\mathrm{rot}}\,\vec{u} - \vec{u}\cdot\overrightarrow{\mathrm{rot}}\,\overrightarrow{M}) + \overrightarrow{\nabla}f\cdot(\vec{u}\wedge\overrightarrow{M}) = \vec{u}\cdot(\overrightarrow{M}\wedge\overrightarrow{\nabla}f - f\,\overrightarrow{\mathrm{rot}}\,\overrightarrow{M})\,.$$

(Démonstration plus rapide : $\oiint_S \epsilon_{ijk}fM_j\,\mathrm{d}S_k = \iiint_V \epsilon_{ijk}(M_j\,\partial_k f + f\partial_k M_j)\,\mathrm{d}^3r$).

▷ Frontières et champs dépendant du temps

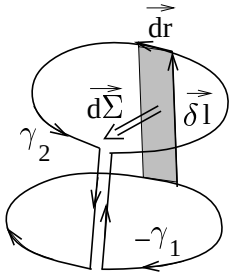


FIGURE 23

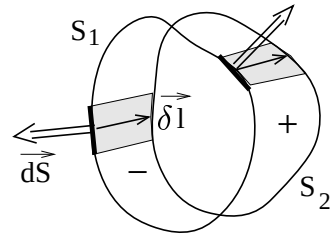


FIGURE 24

Un déplacement infinitésimal $\overrightarrow{\delta l}(\vec{r})$ "continu" des points de γ ou S (fermés) entraîne les variations :

$$\delta\left[\oint_\gamma \overrightarrow{A}\cdot\overrightarrow{\mathrm{d}l}\right] = \oint_\gamma \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}\cdot(\overrightarrow{\delta l}\wedge\overrightarrow{\mathrm{d}r}) \quad ; \quad \delta\left[\iiint_V f\,\mathrm{d}^3r\right] = \oiint_S f\,\overrightarrow{\delta l}\cdot\overrightarrow{\mathrm{d}S}\,.$$

DÉMONSTRATION : $\mathcal{C}_{\gamma_2}(\overrightarrow{A}) - \mathcal{C}_{\gamma_1}(\overrightarrow{A}) = \mathcal{C}_{\{\gamma_2-\gamma_1\}}(\overrightarrow{A})$ où $\{\gamma_2-\gamma_1\}$ est le circuit fermé de la figure 23 qui limite une surface latérale (supposée orientable) de vecteur $\overrightarrow{\mathrm{d}\Sigma} = \overrightarrow{\delta l}\wedge\overrightarrow{\mathrm{d}r}$. La variation de l'intégrale triple a elle simplement pour origine les volumes hachurés (algébriques) $\overrightarrow{\mathrm{d}S}\cdot\overrightarrow{\delta l}$ (figure 24).

Applications en électromagnétisme. Si le champ $\overrightarrow{A}$ et le déplacement $\overrightarrow{\delta l}$ dépendent du temps (avec $\overrightarrow{\delta l} = \vec{v}_\gamma(\vec{r},t)\,\mathrm{d}t$), alors :

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\oint_{\gamma_t}\overrightarrow{A}\cdot\overrightarrow{\mathrm{d}r}\right] = \oint_{\gamma_t}\Big(\frac{\partial\overrightarrow{A}}{\partial t} + \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}\wedge\vec{v}_\gamma\Big)\cdot\overrightarrow{\mathrm{d}r}\ .$$

(Ceci résulte de

$$\mathcal{C}_{\gamma_{t+\mathrm{d}t}}\big(\overrightarrow{A}(\vec{r},t+\mathrm{d}t)\big) - \mathcal{C}_{\gamma_t}\big(\overrightarrow{A}(\vec{r},t)\big) = \mathcal{C}_{\gamma_{t+\mathrm{d}t}}\big(\overrightarrow{A}(\vec{r},t+\mathrm{d}t) - \overrightarrow{A}(\vec{r},t)\big) + \mathcal{C}_{\{\gamma_{t+\mathrm{d}t}-\gamma_t\}}\big(\overrightarrow{A}(\vec{r},t)\big)$$

et du fait que $\mathcal{C}_{\gamma_{t+\mathrm{d}t}}\big(\overrightarrow{A}(\vec{r},t+\mathrm{d}t) - \overrightarrow{A}(\vec{r},t)\big) = \mathcal{C}_{\gamma_t}\Big(\dfrac{\partial\overrightarrow{A}}{\partial t}\Big)\,\mathrm{d}t$ au premier ordre). Une première application est la **loi de Lenz** $e = -\dfrac{\mathrm{d}\varphi}{\mathrm{d}t}$, qui relie la circulation du champ électromoteur $\overrightarrow{E}_{\mathrm{em}} = -\partial_t\overrightarrow{A} + \vec{v}_\gamma\wedge\overrightarrow{B}$ au flux de $\overrightarrow{B} = \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}$ pour des fils conducteurs en mouvement à la vitesse $\vec{v}_\gamma$. Remarque : les contre exemples connus à cette loi sont des cas où, soit $\vec{v}_\gamma$ est discontinue (dynamo unipolaire ; figure 25a), soit γ_t balaye une surface non orientable (galvanomètre à champ radial ; figure 25b).

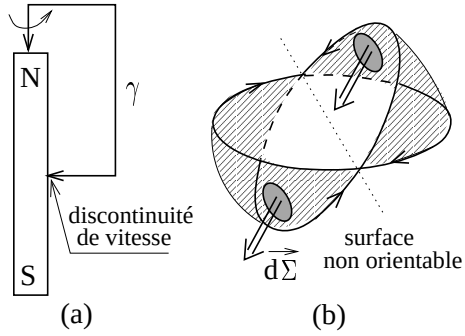


FIGURE 25

Une seconde application est le **bilan énergétique** (pouvant être étendu à plusieurs circuits) $\displaystyle\oint_\gamma I\frac{\partial\overrightarrow{A}}{\partial t}\cdot\overrightarrow{\mathrm{d}r} = I\,\frac{\mathrm{d}\varphi_S(\overrightarrow{B})}{\mathrm{d}t} - \oint_\gamma I(\overrightarrow{\mathrm{d}r}\wedge\overrightarrow{B})\cdot\vec{v}_\gamma$, qui relie la dérivée temporelle de l'énergie magnétique aux puissances fournies par des sources électriques et par des sources mécaniques s'opposant aux forces de Laplace (*cf.* sections 7.3.1 et 7.5.2).

Applications en hydrodynamique. Si dans l'intégrale triple le domaine V_t dépend du temps, chaque point étant entraîné par le champ de vitesse $\vec{v}_{\mathrm{e}}$) (qui n'est pas nécessairement celui du fluide), on a pour toute fonction $f(\vec{r},t)$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left[\iiint_{V_t} f(\vec{r},t)\,\mathrm{d}^3r\right] &= \iiint_{V_t}\frac{\partial f(\vec{r},t)}{\partial t}\,\mathrm{d}^3r + \oiint_{S_t} f(\vec{r},t)\,\vec{v}_{\mathrm{e}}\cdot\overrightarrow{\mathrm{d}S} = \iiint_{V_t}\Big(\frac{\partial f}{\partial t} + \mathrm{div}(f\vec{v}_{\mathrm{e}})\Big)\,\mathrm{d}^3r\\ &= \iiint_{V_t}\Big(\frac{\mathrm{d}f}{\mathrm{d}t} + f\,\mathrm{div}\vec{v}_{\mathrm{e}}\Big)\,\mathrm{d}^3r \quad \text{avec } \frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \vec{v}_{\mathrm{e}}\cdot\overrightarrow{\nabla}\ .\end{aligned}$$

DÉMONSTRATION. La première égalité résulte de ce que la différence des intégrales de $f(\vec{r},t+\mathrm{d}t)$ sur $V_{t+\mathrm{d}t}$ et de $f(\vec{r},t)$ sur V_t est la somme de l'intégrale de $f(\vec{r},t+dt)-f(\vec{r},t)$

sur $V_{t+\mathrm{d}t}$, ou V_t au premier ordre, et de l'intégrale de $f(\vec{r},t)$ sur $\{V_{t+\mathrm{d}t}-V_t\}$. Les autres sont des conséquences immédiates. La dernière expression se déduit aussi de $\dfrac{\mathrm{d}}{\mathrm{d}t}\left(f\,\mathrm{d}^3r\right)$ avec $\dfrac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{d}^3r = \mathrm{div}\vec{v}_{\mathrm{e}}\,\mathrm{d}^3r$ (ou justifie cette égalité si on choisit $f=1$).

Si $f=\rho$ (masse volumique) et $\vec{v}_{\mathrm{e}}=\vec{v}$ (vitesse du fluide), la **conservation de la masse** dans le volume V_t entraîné par un fluide se traduit par :

$$\frac{\partial\rho}{\partial t}+\mathrm{div}(\rho\vec{v})=0 \qquad (\text{ou } \frac{\mathrm{d}\rho}{\mathrm{d}t}+\rho\,\mathrm{div}\vec{v}=0) \quad \text{L.C. de la masse .}$$

Plus généralement si G est une grandeur de densité volumique $\rho_{_G}=\rho g$ (g **grandeur massique**, $g=1$ pour la masse, $g=V=\rho^{-1}$ pour le volume, $\vec{g}=\vec{v}$ pour la quantité de mouvement, $g=\dfrac{1}{2}\,v^2$ pour l'énergie cinétique et $g=e$ pour l'énergie interne), on a :

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[G_{V_t}=\iiint_{V_t}\rho g\,\mathrm{d}^3r\right]=\iiint_{V_t}\Big(\frac{\partial(\rho g)}{\partial t}+\mathrm{div}(\rho g\vec{v})\Big)\,\mathrm{d}^3r=\iiint_{V_t}\rho\,\frac{\mathrm{d}g}{\mathrm{d}t}\,\mathrm{d}^3r\ .$$

La dernière égalité résulte de la L.C. de la masse et s'obtient aussi directement en écrivant que $\mathrm{d}_t(\rho g\,\mathrm{d}^3r)=\rho\,\mathrm{d}^3r\,\mathrm{d}_t g$ car $\mathrm{d}m=\rho\,\mathrm{d}^3r$ est constant.

7.2.4 Bilans de grandeurs ; applications aux milieux continus

▷ Circulation automobile

Cet exemple simple sert d'introduction aux divers raisonnements sur les bilans. Sur une voie de circulation, soient $n(x,t)$ la densité linéique de voitures se déplaçant à la vitesse $v(x,t)$ et $j(x,t)=n(x,t)\,v(x,t)$ leur flux. La L.C. des voitures se traduit, pour leur nombre N_{AB} entre deux points A et B, par le bilan $\dfrac{\mathrm{d}N_{AB}}{\mathrm{d}t}=j_{_A}-j_{_B}$ si A et B sont fixes, et par le bilan $\dfrac{\mathrm{d}N_{AB}}{\mathrm{d}t}=0$ si A et B sont entraînés avec les voitures. Pour A et B "proches" ($x_{_A}=x$, $x_{_B}=x+\mathrm{d}x$) et n et j "continus", le premier s'écrit

$$\frac{\mathrm{d}}{\mathrm{d}t}(n\,\mathrm{d}x)=\frac{\partial n}{\partial t}\,\mathrm{d}x=j(x,t)-j(x+\mathrm{d}x,t)\ \text{ soit }\ \partial_t n+\partial_x j=0\ ,$$

et le second

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(n(x_{_B}-x_{_A})\big)=\frac{\mathrm{d}n}{\mathrm{d}t}\,\mathrm{d}x+n\,\mathrm{d}v=0\ \text{ soit }\ \mathrm{d}_t n+n\,\partial_x v=0$$

(résultats équivalents).

Si on s'intéresse à une grandeur g transportée par chaque voiture ($g=m,\ mv,\ \dfrac{1}{2}\,mv^2\cdots$), sa densité est $\rho_{_G}=ng$ et son flux (ou courant) est $j_{_G}=ngv$. La quantité

$$\tau=\partial_t(ng)+\partial_x(ngv)=n\,\partial_t g+nv\,\partial_x g\ \text{(L.C. des voitures)}\ =n\,\mathrm{d}_t g$$

est l'apport (par unité de longueur) aux voitures de la grandeur considérée (force linéique exercée par la route si $g=mv$).

S'il existe une **onde de choc** (provoquée par des voitures arrétées à un feu par exemple), alors pour A et B fixes situés de part et d'autre de la discontinuité qui se déplace à la vitesse V, le bilan est $\mathrm{d}N_{AB} = (n_1 - n_2)V\,\mathrm{d}t$ (car dans l'intervalle de temps $\mathrm{d}t$ une densité linéique n_2 est remplacée par n_1 sur la longueur $V\,\mathrm{d}t$; figure 26).

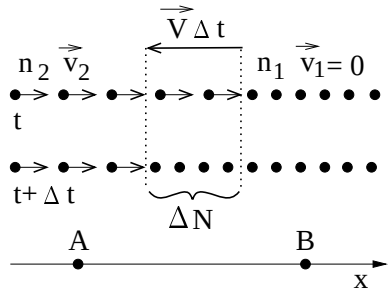


FIGURE 26

De

$$\frac{\mathrm{d}N_{AB}}{\mathrm{d}t} = (n_1 - n_2)V = j_{_A} - j_{_B} = n_1 v_1 - n_2 v_2 \ ,$$

on déduit $V = \dfrac{n_1 v_1 - n_2 v_2}{n_1 - n_2}$. On notera enfin qu'une distribution continue peut induire une onde de choc si il existe une relation (phénoménologique) $j(n)$ entre densité et courant (flux) ; en effet la L.C. qui s'écrit alors $\partial_t n + c(n)\,\partial_x n = 0$ avec $c(n) = \dfrac{\mathrm{d}j}{\mathrm{d}n}$ est l'équation d'onde discutée à la section 7.1.1.

▷ Ecritures générales d'un bilan

Soit G_V une grandeur extensive relative à un **volume fixe**. On peut souvent distinguer dans $\dfrac{\mathrm{d}G_V}{\mathrm{d}t}$ ce qui est dû à des échanges à travers la frontière S (notion de flux associé à une densité de courant $\vec{j}_{_G}$), de la "production sur place" de grandeur (associée à des termes de source algébriques $\tau_{_G}$). Le bilan s'écrit, sous forme intégrale

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[G_V = \iiint_V \rho_{_G}\,\mathrm{d}^3 r\right] = \iiint_V \frac{\partial \rho_{_G}}{\partial t}\,\mathrm{d}^3 r = -\oiint_S \vec{j}_{_G}\cdot\overrightarrow{\mathrm{d}S} + \iiint_V \tau_{_G}\,\mathrm{d}^3 r \ ,$$

et sous forme différentielle :

$$\frac{\partial \rho_{_G}}{\partial t} + \mathrm{div}\vec{j}_{_G} = \tau_{_G} \ .$$

Si V_t est un **volume dépendant du temps**, la comparaison de $\dfrac{\mathrm{d}G_V}{\mathrm{d}t}$ et de $\dfrac{\mathrm{d}G_{V_t}}{\mathrm{d}t}$ montre qu'il suffit de remplacer $\vec{j}_{_G}$ par $\vec{j}_{_G}{}' = \vec{j}_{_G} - \rho_{_G}\vec{v}_{_S}$ ou $\vec{v}_{_S}$ est la vitesse d'un point de la surface ($\vec{v}_{\mathrm{S}} = \vec{v}$ si V_t suit le fluide).

▷ Lois de conservation (L.C.)

La grandeur G est dite **conservée localement** si $\tau_{_G} = 0$ (pas de création). C'est le cas de la masse, de la charge électrique, de la quantité de mouvement en absence de forces à distance... Par contre $\tau_S \geq 0$ pour l'entropie (*cf.* section 8.2.3).

REMARQUES. 1) Notamment dans le **cas de l'énergie** il est important de bien préciser la "forme" considérée : énergie cinétique, ou interne, ou relative à un sous système (par exemple à des particules dans un champ électromagnétique), ou à une onde (qui peut être absorbée par le milieu), etc ; en effet même si $\tau_{_G} = 0$ pour la grandeur totale, en général $\tau_{_G} \neq 0$ pour chaque "forme". 2) Si S est une surface de discontinuité (**onde de choc**), la L.C. s'écrit (égalité des flux de part et d'autre de la discontinuité) :

$$(\vec{j}_{G,2} - \rho_{G,2}\vec{v}_{_S}) \cdot \overrightarrow{\mathrm{d}S} = (\vec{j}_{G,1} - \rho_{G,1}\vec{v}_{_S}) \cdot \overrightarrow{\mathrm{d}S} \qquad (v_{_S} = \frac{j_{G,2} - j_{G,1}}{\rho_{G,2} - \rho_{G,1}} \text{ à une dimension}) \ .$$

▷ Lois de conservation microscopiques

Les L.C. locales définies ci-dessus concernent des grandeurs macroscopiques. Il existe aussi en physique des L.C. microscopiques : pour des particules ponctuelles i possèdant une caractéristique g_i constante au cours du temps (nombre $g_i \equiv 1$, masse m_i, charge q_i...), et décrivant des trajectoires $\vec{r}_i(t)$ à la vitesse $\vec{v}_i(t)$, on a automatiquement :

$$\partial_t \rho_{_G}^{\text{micro}} + \mathrm{div}\vec{j}_{_G}^{\text{micro}} = 0 \text{ avec } \rho_{_G}^{\text{micro}} = \sum_i g_i\, \delta\big(\vec{r} - \vec{r}_i(t)\big) \text{ et } \vec{j}_{_G}^{\text{micro}} = \sum_i g_i\, \vec{v}_i(t)\, \delta\big(\vec{r} - \vec{r}_i(t)\big).$$

Cette loi reste valable si on remplace $\delta(\vec{r})$ par une fonction "lissante" $f(\vec{r})$.
DÉMONSTRATION : pour tout f on a $\partial_t f\big(\vec{r} - \vec{r}_i(t)\big) = -\vec{v}_i(t) \cdot \overrightarrow{\nabla} f = -\mathrm{div}\big(f\vec{v}_i(t)\big)$; si g_i dépendait du temps on aurait un terme de source $\tau_{_G}^{\text{micro}} = \sum_i \dot{g}_i(t)\, f\big(\vec{r} - \vec{r}_i(t)\big)$. C'est ainsi, par des opérations de **lissage** (moyenne), que beaucoup de bilans macroscopiques peuvent être déduits de bilans microscopiques.

▷ Flux en hydrodynamique et élasticité

Quand on suit dans son mouvement avec le milieu un élément de surface $\overrightarrow{\mathrm{d}S}$ (1 et 2 désignant le milieu de part et d'autre), le flux de masse, le flux de quantité de mouvement (force de contact exercée par 1 sur 2) et le flux d'énergie ("puissance de la force de contact + flux thermique") valent respectivement :

$$\mathrm{d}\varphi'_m = 0 \quad ; \quad \mathrm{d}\vec{\varphi}'_{\vec{p}} = [\mathbf{P}]\, \overrightarrow{\mathrm{d}S} \ \ (\text{ou } -[\boldsymbol{\sigma}]\, \overrightarrow{\mathrm{d}S}) \quad ; \quad \mathrm{d}\varphi'_E = \vec{v} \cdot [\mathbf{P}]\, \overrightarrow{\mathrm{d}S} + \vec{j}_{\text{th}} \cdot \overrightarrow{\mathrm{d}S} \ .$$

Si $\overrightarrow{\mathrm{d}S}$ est fixe, comme $\vec{j}_{_G} = \vec{j}'_{_G} + \rho_{_G}\, \vec{v}$, il faut ajouter à ces flux $\rho_{_G}\, \vec{v} \cdot \overrightarrow{\mathrm{d}S}$ avec $\rho_{_G} = \rho$ ou $\rho\vec{v}$ ou $\rho(\frac{1}{2}\, v^2 + e)$. Pour un **fluide parfait** ($[\mathbf{P}] \equiv P$ pression, $\vec{j}_{\text{th}} = 0$) on obtient (*cf.* aussi la section 3.2.4) :

$$\mathrm{d}\varphi_m = \rho\vec{v} \cdot \overrightarrow{\mathrm{d}S} \quad ; \quad \mathrm{d}\vec{\varphi}_{\vec{p}} = \rho\vec{v}(\vec{v} \cdot \overrightarrow{\mathrm{d}S}) + P\, \overrightarrow{\mathrm{d}S} \quad ; \quad \mathrm{d}\varphi_E = \Big(\frac{1}{2}\, v^2 + e + \frac{P}{\rho}\Big)\, \rho\vec{v} \cdot \overrightarrow{\mathrm{d}S} \ .$$

Application : **vitesse d'une onde sonore** à une dimension. Les L.C. de la masse et de la quantité de mouvement en présence d'une discontinuité donnent :

$$v_{_S} = \frac{\rho_2 v_2 - \rho_1 v_1}{\rho_2 - \rho_1} = \frac{(P_2 + \rho_2 v_2^2) - (P_1 + \rho_1 v_1^2)}{\rho_2 v_2 - \rho_1 v_1} \ .$$

Si la discontinuité est faible et le fluide proche du repos ($\rho_i v_i^2 \ll P_i$), alors $v_{_S}^2 \simeq \dfrac{\mathrm{d}P}{\mathrm{d}\rho}$.

REMARQUE. On peut ajouter à $\mathrm{d}\varphi_E$ le terme $\rho g z\,\vec{v}\cdot\overrightarrow{\mathrm{d}S}$ si l'énergie potentielle de pesanteur n'est pas prise en compte par ailleurs (*cf.* bilans locaux ci-dessous). Les bilans déduits des flux ci-dessus recouvrent une grande quantité d'applications pratiques. Beaucoup concernent la quantité $\frac{1}{2}\,v^2+e+\frac{P}{\rho}+gz$ qui, en écoulement permanent, reste constante le long d'une ligne de courant si l'énergie du fluide est conservée (**théorème de Bernouilli**), ou varie d'une quantité reliée aux travaux et échanges thermiques effectués par une masse unité de fluide lors de son écoulement entre les points considérés. On trouvera des exemples dans les livres d'hydrodynamique pour ingénieurs.

▷ Bilans locaux et équations des milieux continus

Pour une masse $\mathrm{d}m$ de matière qui occupe un volume élémentaire d^3r et se déplace à la vitesse $\vec{v}$, les bilans de quantité de mouvement et d'énergie s'écrivent (l'effet des forces de contraintes ayant été décrit à la section 7.2.3) :

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(\mathrm{d}m\,\vec{v}) = \vec{f_c}\,\mathrm{d}^3r + \mathrm{d}m\,\vec{g} + \vec{f_a}\,\mathrm{d}^3r = \mathrm{d}m\,\frac{\mathrm{d}\vec{v}}{\mathrm{d}t}\ ,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{d}m\left(\frac{\vec{v}^2}{2}+e\right)\right) = \mathrm{div}\big([\boldsymbol{\sigma}]\,\vec{v}\big)\,\mathrm{d}^3r - \mathrm{div}\vec{j}_{\mathrm{th}}\,\mathrm{d}^3r + \vec{v}\cdot(\mathrm{d}m\,\vec{g}+\vec{f_a}\,\mathrm{d}^3r) + \tau_a\,\mathrm{d}^3r\ .$$

$\mathrm{d}m\,\vec{g}$ est la force due au champ de pesanteur local ; $\vec{f_a}$ désigne des forces volumiques autres, par exemple de Laplace $\vec{j}\wedge\overrightarrow{B}$ en magnétohydrodynamique, de Coriolis $-2\rho\vec{\omega}\wedge\vec{v}$ et "centrifuge" dans un référentiel non inertiel, *etc.* ; τ_a désigne les sources d'énergie autres (effet Joule par exemple). En divisant les deux membres par d^3r, le bilan de quantité de mouvement s'écrit :

$$\rho\,\frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = \rho\left(\frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\overrightarrow{\nabla})\,\vec{v}\right) = \vec{f_c}+\rho\vec{g}+\vec{f_a}\ .$$

On en déduit, en faisant le produit scalaire avec $\vec{v}$, le bilan d'énergie cinétique et potentielle :

$$\rho\,\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\vec{v}^2+gz\right) = (\vec{f_c}+\vec{f_a})\cdot\vec{v} \quad \text{car } \rho\vec{g}\cdot\vec{v} = -\rho\,\frac{\mathrm{d}(gz)}{\mathrm{d}t}\ .$$

Il permet de récrire le bilan d'énergie sous la forme, ne faisant intervenir que l'énergie interne (**premier principe**) :

$$\rho\,\frac{\mathrm{d}e}{\mathrm{d}t} = \mathrm{tr}\big([\boldsymbol{\sigma}][\dot{\boldsymbol{\epsilon}}]\big) - \mathrm{div}\vec{j}_{\mathrm{th}} + \tau_a \qquad ([\dot{\boldsymbol{\epsilon}}]_{ij} = \frac{1}{2}\,(\partial_i v_j + \partial_j v_i))\ .$$

Dans la suite $\vec{j}_{\mathrm{th}}=0$, $\vec{f_a}=0$ et $\tau_a=0$. Par exemple pour un **fluide parfait** ($\vec{f_c} = -\overrightarrow{\mathrm{grad}}\,P$), on obtient l'**E.D.P d'Euler** $\rho\,\dfrac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\overrightarrow{\mathrm{grad}}\,P+\rho\,\vec{g}$ qui, si ρ est constant (ou si P ne dépend que de ρ), fait que $\dfrac{\mathrm{d}\vec{v}}{\mathrm{d}t}$ est un gradient et conduit à tous les théorèmes vus à la section 7.1.3 ; la puissance interne volumique $\mathrm{tr}\big([\boldsymbol{\sigma}][\dot{\boldsymbol{\epsilon}}]\big) = -P\,\mathrm{div}\vec{v}$ correspond au terme $-P\,\dfrac{\mathrm{d}V}{\mathrm{d}t}$ en thermodynamique.

AUTRES EXEMPLES. 1) **Fluide visqueux**. Il faut alors ajouter la force de composantes $f_{c,i} = \sum_j \partial_j\sigma_{ij}$ avec, dans le cas linéaire (*cf.* section 4.3.1), $\sigma_{ij} = \eta(\partial_i v_j+\partial_j v_i)+\delta_{ij}\left(\zeta-\frac{2\eta}{3}\right)\mathrm{div}\vec{v}$; ceci donne

$$\rho\,\tfrac{\mathrm{d}\vec{v}}{\mathrm{d}t} = -\overrightarrow{\mathrm{grad}}\,P+\rho\vec{g}+\eta\,\Delta\vec{v}+(\zeta+\tfrac{\eta}{3})\,\overrightarrow{\mathrm{grad}}\,(\mathrm{div}\vec{v})$$

qui, pour un fluide incompressible ($\mathrm{div}\vec{v} = 0$), est l'**E.D.P. de Navier-Stokes**. On vérifie que $\mathrm{tr}\,([\boldsymbol{\sigma}][\dot{\boldsymbol{\epsilon}}]) > 0$ (augmentation locale de l'énergie interne du fluide). 2) **Milieu élastique linéaire**. Avec $\overrightarrow{\psi}$ désignant le champ de déplacement (infinitésimal) et $\sigma_{ij} = \mu\,(\partial_i\psi_j + \partial_j\psi_i) + \delta_{ij}\,\left(\chi^{-1} - \frac{2\mu}{3}\right)\mathrm{div}\,\overrightarrow{\psi}$, on obtient l'**E.D.P. des ondes élastiques** :

$$\rho\,\frac{\partial^2\overrightarrow{\psi}}{\partial t^2} = \mu\,\Delta\overrightarrow{\psi} + \left(\chi^{-1} + \frac{\mu}{3}\right)\overrightarrow{\mathrm{grad}}\,(\mathrm{div}\,\overrightarrow{\psi})\ .$$

$\mathrm{tr}\,([\boldsymbol{\sigma}][\dot{\boldsymbol{\epsilon}}])$ est alors la dérivée temporelle d'une énergie élastique.

7.3 APPLICATIONS À LA MÉCANIQUE ET À L'OPTIQUE GÉOMÉTRIQUE

7.3.1 Mécanique et fonctions énergie potentielle

Dans de nombreux cas importants, les forces (et les couples) peuvent se calculer à partir de la différentielle d'une fonction des positions (et orientations) des corps concernés appelée énergie potentielle E_p. Cette fonction dépend aussi de paramètres décrivant l'état des corps ou leur environnement, paramètres qui, par définition, restent fixés dans la différentiation. Sa détermination à partir des énergies connues en physique dépend de façon cruciale du choix effectué pour les paramètres.

▷ Système de points matériels

Pour des points matériels (isotropes) l'énergie potentielle est une fonction $E_p(\vec{r}_1, \vec{r}_2 \cdots \vec{r}_n)$ de leurs positions telle que les forces $\vec{f}_i$ sur chacun d'eux s'en déduisent par :

$$\mathrm{d}E_p \overset{\mathrm{def}}{=} -\sum_i \vec{f}_i \cdot \overrightarrow{\mathrm{d}r}_i \quad \text{soit} \quad \vec{f}_i = -\overrightarrow{\nabla}_i\,E_p \ \text{ avec } \ \overrightarrow{\nabla}_i \equiv \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial y_i}, \frac{\partial}{\partial z_i}\right)\ .$$

L'équilibre ($\vec{f}_i = 0$) correspond à E_p extremum ; il est stable si $\delta^2 E_p > 0$.

La **force résultante** $\overrightarrow{F}_\alpha = \sum_{i\in\alpha} \vec{f}_i$ sur un sous ensemble α de points s'obtient en remarquant que $\mathrm{d}E_p = -\overrightarrow{F}_\alpha \cdot \overrightarrow{\mathrm{d}r}_\alpha$ lorsque $\overrightarrow{\mathrm{d}r}_{i\in\alpha} = \overrightarrow{\mathrm{d}r}_\alpha$ et $\overrightarrow{\mathrm{d}r}_{i\notin\alpha} = 0$ (donc en considérant par la pensée une translation rigide de l'ensemble α, les autres points restant fixes). Pour le **moment résultant** des forces $\overrightarrow{\Gamma}_\alpha = \sum_{i\in\alpha} \vec{r}_i \wedge \vec{f}_i$, on remarque aussi que : $\mathrm{d}E_p = -\overrightarrow{\Gamma}_\alpha \cdot \overrightarrow{\mathrm{d}\theta}_\alpha$ lorsque $\overrightarrow{\mathrm{d}r}_{i\in\alpha} = \overrightarrow{\mathrm{d}\theta}_\alpha \wedge \vec{r}_i$ et $\overrightarrow{\mathrm{d}r}_{i\notin\alpha} = 0$ (rotation rigide de l'ensemble α, les autres points restant fixes). Si E_p est invariante par translation (resp. rotation) la force résultante totale (resp. le moment total) est nulle.

REMARQUE. Le travail élémentaire des forces $\vec{f}_i$ étant $\delta W = \sum_i \overrightarrow{f}_i \cdot \overrightarrow{\mathrm{d}r}_i$ la différentielle

$$\mathrm{d}E_p = -\delta W$$

est le travail δW_{sm} qu'une **source mécanique** (un "opérateur"), exerçant les forces $-\vec{f}_i$, devrait effectuer pour rendre les déplacements quasi-statiques.

EXEMPLE 1. **Electrostatique** : $E_p = (4\pi\epsilon_0)^{-1}\sum_{(i,j)} r_{ij}^{-1} q_i q_j$. Divisons les charges en deux groupes α et β. Comme les interactions ne dépendent que de $r_{ij} = |\vec{r}_i - \vec{r}_j|$ il suffit, pour avoir les actions de β sur α (ou de α sur β), de considérer l'**énergie poten-**

tielle d'interaction qui est la restriction de la double somme aux couples ($i\in\alpha$, $j\in\beta$) :

$$E_p^{int} = \sum_{i\in\alpha} q_i V_{ext}(\vec{r}_i) \quad \text{avec} \quad V_{ext}(\vec{r}_i) = (4\pi\epsilon_0)^{-1} \sum_{j\in\beta} r_{ij}^{-1} q_j \ .$$

Par exemple si α est un **dipole** $\vec{p} = \sum_i q_i \vec{r}_i$ ($\sum_i q_i = 0$) situé en $\vec{r}_\alpha \simeq \vec{r}_i$, on a $E_p^{int} = \sum_i q_i V_{ext}(\vec{r}_i) = -\vec{p}\cdot\overrightarrow{E}(\vec{r}_\alpha)$ avec $\overrightarrow{E} = -\overrightarrow{\nabla} V_{ext}$. Dans une translation infinitésimale selon x, $\mathrm{d}E_p^{int} = -(p_x\,\partial_x E_x + p_y\,\partial_x E_y + p_z\,\partial_x E_z)\,\mathrm{d}x$ d'où $F_x = (\vec{p}\cdot\overrightarrow{\nabla})\,E_x$ (en utilisant $\overrightarrow{\mathrm{rot}}\,\overrightarrow{E} = 0$) et $\overrightarrow{F} = (\vec{p}\cdot\overrightarrow{\nabla})\,\overrightarrow{E}$. Dans une rotation infinitésimale centrée en $\vec{r}_\alpha$, on a $\mathrm{d}E_p^{int} = -(\mathrm{d}\vec{\theta}\wedge\vec{p})\cdot\overrightarrow{E}$ d'où le couple $\overrightarrow{C} = \vec{p}\wedge\overrightarrow{E}$. (Avec une rotation de centre O on obtiendrait le moment $\overrightarrow{\Gamma} = \overrightarrow{C} + \vec{r}_\alpha\wedge\overrightarrow{F}$.) Ces résultats ne dépendent pas du caractère déformable ou indéformable (induit ou pas) du dipole. Si β est aussi un dipole, l'invariance par translation et par rotation de E_p entraîne $\overrightarrow{F}_{\beta\to\alpha} + \overrightarrow{F}_{\alpha\to\beta} = 0$ et $\overrightarrow{C}_{\beta\to\alpha} + \overrightarrow{C}_{\alpha\to\beta} + (\vec{r}_\beta - \vec{r}_\alpha)\wedge\overrightarrow{F}_{\alpha\to\beta} = 0$; en général $\overrightarrow{F}_{\alpha\to\beta} = -\overrightarrow{F}_{\beta\to\alpha}$ n'est pas parallèle à $\vec{r}_\beta - \vec{r}_\alpha$ car un dipole, même ponctuel, n'est pas isotrope.

Remarque : ci-dessus E_p est une somme de fonctions de $\vec{r}_i - \vec{r}_j$. Ceci n'est pas toujours le cas, par exemple pour les atomes d'une molécule ou d'un cristal avec des liaisons délocalisées ; il n'y a alors aucun moyen naturel de décomposer $\vec{f}_i = -\overrightarrow{\nabla}_i E_p$ sous la forme traditionnelle $\vec{f}_i = \sum_{j\neq i} \vec{f}_{j\to i}$.

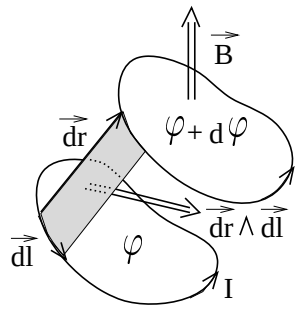


FIGURE 27

EXEMPLE 2. **Circuit électrique dans un champ extérieur** $\overrightarrow{B}$ (figure 27).
Dans un déplacement infinitésimal "virtuel" $\overrightarrow{\mathrm{d}r}$ de chaque point $\vec{r}$ du circuit, le travail des forces de Laplace $I\,\overrightarrow{\mathrm{d}l}\wedge\overrightarrow{B}$ s'écrit $\delta W = \oint(I\,\overrightarrow{\mathrm{d}l}\wedge\overrightarrow{B})\cdot\overrightarrow{\mathrm{d}r} = \oint(I\,\overrightarrow{\mathrm{d}r}\wedge\overrightarrow{\mathrm{d}l})\cdot\overrightarrow{B} = I\,\mathrm{d}\varphi$ (conservation du flux de $\overrightarrow{B}$). La force et le moment totaux qu'exerce $\overrightarrow{B}$ sur le circuit s'obtiennent donc à partir de $E_p = -I\varphi$ (où φ dépend des 6 coordonnées qui déterminent la position et l'orientation dans l'espace du circuit rigide). Pour une petite spire, $E_p = -\vec{\mu}\cdot\overrightarrow{B}$ (avec $\vec{\mu} = I\,\overrightarrow{S}$) comme pour un dipole électrique.

▷ Fonction E_p et énergie cinétique

E_p a un lien mathématique simple avec l'énergie cinétique E_c (relativiste ou non). Comme $\mathrm{d}E_p \overset{\text{def}}{=} -\sum_i \vec{f}_i\cdot\overrightarrow{\mathrm{d}r}_i$ (à paramètres fixes) et $\dfrac{\mathrm{d}E_c}{\mathrm{d}t} = \sum_i \vec{f}_i\cdot\vec{v}_i$ on a, si les paramètres dépendent du temps :

$$\frac{\mathrm{d}}{\mathrm{d}t}(E_p + E_c) = \frac{\partial E_p}{\partial t} \ .$$

La loi $E_p + E_c = \mathrm{C^{ste}}$ pour un système isolé a beaucoup d'applications en mécanique.

Théorème du viriel. Toute fonction homogène E_p d'ordre n vérifie :

$nE_p = \sum_i \vec{r}_i \cdot \overrightarrow{\nabla_i} E_p$ (E.D.P d'Euler) $= -\sum_i \vec{r}_i \cdot \vec{f}_i$ (viriel des forces) $= -\frac{\mathrm{d}}{\mathrm{d}t}(\sum_i \vec{r}_i \cdot m_i \vec{v}_i) + 2E_c$.

Si le domaine de variation des positions et des vitesses est limité, la moyenne sur un grand intervalle de temps $< \frac{\mathrm{d}}{\mathrm{d}t}(\sum_i \vec{r}_i \cdot m_i \vec{v}_i) >$ est nulle et :

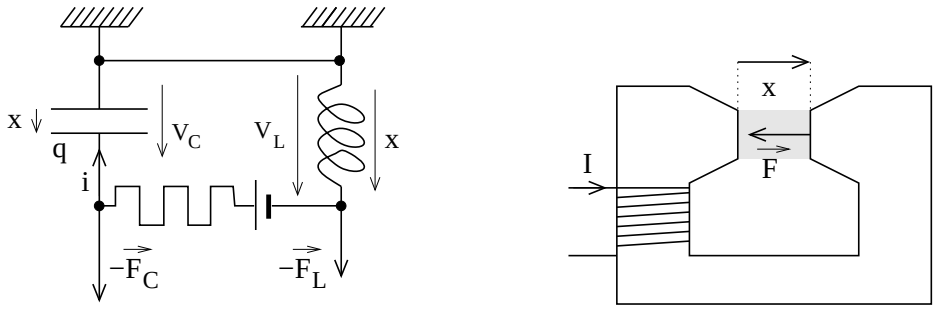
$$2 < E_c >= n < E_p > \ .$$

Les applications principales correspondent à $n = -1$ (interactions gravitationnelles en astrophysique ou électrostatiques pour l'atome de Bohr) et $n = 2$ (oscillateurs harmoniques couplés).

▷ Fonctions E_p en électromagnétisme

Partons d'un exemple élémentaire de bilan d'énergie où la connaissance (physique) des énergies impliquées et des puissances fournies par les générateurs électriques permet de déduire les travaux élémentaires δW_{sm} des sources mécaniques.

Circuit RLC avec inductance et capacité déformables (figure 28a).



(a) (b)

FIGURE 28

De $I = \dfrac{\mathrm{d}q}{\mathrm{d}t}$, $V_L = \dfrac{\mathrm{d}\varphi}{\mathrm{d}t}$ (loi de Lenz avec $\varphi = LI$) et $V_C = \dfrac{q}{C}$, on déduit les égalités (qui concernent les énergies électromagnétiques $\dfrac{1}{2}LI^2$ et $\dfrac{1}{2}CV_C^2$) :

$$\mathrm{d}\Big(\frac{1}{2}LI^2\Big) = I\,\mathrm{d}\varphi - \frac{1}{2}I^2\,\mathrm{d}L \quad ; \quad \mathrm{d}\Big(\frac{1}{2}CV_C^2\Big) = V_C\,\mathrm{d}q - \frac{1}{2}V_C^2\,\mathrm{d}C \ .$$

Chaque différentielle apparaît comme la somme $\mathrm{d}E_{em} = \delta W_{se} + \delta W_{sm}$ de travaux (connus) fournis par des sources électriques ($V_L\,\mathrm{d}q$ et $V_C\,\mathrm{d}q$), et de travaux (liés aux déformations) qu'il est naturel d'attribuer à des sources mécaniques ($\delta W_{sm} = -\dfrac{1}{2}I^2\,\mathrm{d}L$ et $\delta W_{sm} = -\dfrac{1}{2}V_C^2\,\mathrm{d}C$). On déduit de $\delta W_{sm} = \mathrm{d}E_p$ que les fonctions E_p sont, au choix, $-\dfrac{1}{2}LI^2$ et $-\dfrac{1}{2}CV_C^2$ (c.a.d. $-E_{em}$) si I et V_C sont les paramètres fixes, ou $\dfrac{1}{2}\dfrac{\varphi^2}{L}$ et $\dfrac{1}{2}\dfrac{q^2}{C}$ (c.a.d. E_{em}) si φ et q sont les paramètres fixes. Pour un solénoïde de longueur x constitué de N spires de section S ($L = \mu_0\dfrac{N^2S}{x}$), la force ("contractante") est : $F_L = \dfrac{1}{2}\dfrac{\mathrm{d}L}{\mathrm{d}x}I^2 =$

$-\frac{1}{2}\,\mu_0\,\frac{N^2 S}{x^2}\,I^2$. De même $F_C = -\frac{1}{2}\,\epsilon_0\,\frac{S}{x^2}\,V_C^2$ pour un condensateur plan avec $C = \epsilon_0\,\frac{S}{x}$.

AUTRES EXEMPLES. Pour un **électroaimant** (figure 28b) dont l'entrefer x a une section S, on a $\mathrm{d}E_{em} = \frac{B^2}{2\mu_0}\,S\,\mathrm{d}x$ si $\varphi = BS$ est constant ($\frac{B^2}{2\mu_0}$ énergie volumique dans l'entrefer) ; la force d'attraction entre les faces vaut $F = -\frac{B^2}{2\mu_0}\,S$. Pour un **ensemble de circuits électriques**, $E_{em} = \frac{1}{2}\sum_i I_i\varphi_i$ et $\delta W_{se} = \sum_i I_i\,\mathrm{d}\varphi_i$ (*cf.* section 7.5.2), $\varphi_i = \sum_j M_{ij} I_j$; on en déduit $\delta W_{sm} = \frac{1}{2}\sum_{ij}\mathrm{d}M_{ij}\,I_i I_j$ et $E_p = -\frac{1}{2}\sum_{ij} M_{ij}\,I_i I_j$ (les courants étant pris comme paramètres). Si on s'intéresse aux seules actions sur le circuit 1, on ne garde que les termes $i = 1, j \neq i$ et $i \neq j, j = 1$ et on retrouve $E_p^{(1)} = -I_1\,\varphi$ où φ est le flux du champ extérieur. Le cas des **conducteurs métalliques** en équilibre se déduit de l'exemple précédent en faisant les changements $I_i \to V_i$ et $\varphi_i \to Q_i$ (pour les variables intensives et extensives) et $M_{ij} \to C_{ij}$.

7.3.2 Optique géométrique : rayons ; surfaces d'onde ; caustiques et problèmes d'extremum

On se place dans l'approximation de l'optique géométrique (où les ondes sont localement planes) et on considère des milieux isotropes. Surfaces d'ondes et rayons sont alors deux concepts reliés par la notion de gradient.

▷ Des rayons aux ondes

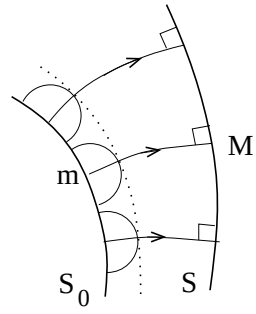


FIGURE 29

Rappelons (*cf.*section 3.2.3) que le chemin optique $L(A,B) \equiv \mathcal{L}(\vec{r}_A, \vec{r}_B)$ calculé le long du rayon joignant deux points A et B correspond à un extremum de $\int_A^B n(\vec{r})\,\mathrm{d}l$, et qu'il a pour différentielle

$$\mathrm{d}\mathcal{L} = n_B\,\hat{u}_B \cdot \overrightarrow{\mathrm{d}r}_B - n_A\,\hat{u}_A \cdot \overrightarrow{\mathrm{d}r}_A$$

($\hat{u}(\vec{r})$ vecteur unitaire porté par le rayon). Les **surfaces d'onde** $L(\vec{r}) =$ constante sont une façon de regrouper les rayons par familles de sorte que, en tout point, $n(\vec{r})\,\hat{u}(\vec{r})$ soit le gradient de $L(\vec{r})$. Par exemple (figure 29) considérons la famille des rayons orthogonaux à une surface S_0 et, pour tout point $M(\vec{r})$ situé sur un rayon issu de $m \in S_0$, posons $L(\vec{r}) = \mathcal{L}(\vec{r}_m, \vec{r})$. La surface S définie par $L(\vec{r}) = L$ (constante) est la surface d'onde déduite de S_0 ($L(\vec{r}) = 0$) après le temps de propagation $\frac{L}{c}$. La surface $L(\vec{r}) = L + \mathrm{d}L$ s'en

déduit en portant sur chaque normale une longueur $\dfrac{\mathrm{d}L}{n(\vec{r})}$; on passe ainsi de proche en proche de S_0 à S. Ceci n'est autre que la **construction de Huygens** des surfaces d'onde comme enveloppes d'ondelettes sphériques, et des rayons comme trajectoires orthogonales à ces surfaces. Elle fournit la solution de l'**E.D.P. eikonale** $|\overrightarrow{\nabla} L(\vec{r})| = n(\vec{r})$ avec la condition aux limites $L =$ constante sur une surface donnée.

REMARQUE. La figure 30 montre, sur le cas particulier $n = \mathrm{C}^{\mathrm{ste}}$ où les rayons sont des droites, que les surfaces S, déduites de S_0 en portant sur les rayons une longueur L, deviennent en général singulières lorsque L croît (*cf.* surface S_2). La singularité se produit sur les **caustiques**, surfaces enveloppe des rayons, et est liée au fait que $L(\vec{r})$ n'est plus définie univoquement (*cf.* exemples ci-dessous).

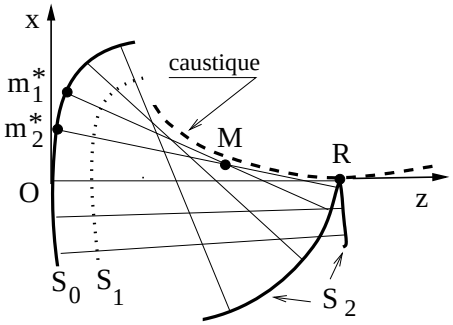


FIGURE 30

▷ Des ondes aux rayons

Soit $\psi(t,\vec{r}) = A(\vec{r})\,\exp -i\big(\omega t - \varphi(\vec{r})\big)$ une onde monochromatique, et soit (par référence à une onde plane $A\exp -i(\omega t - \vec{k}\,.\,\vec{r})$)

$$\vec{k}(\vec{r}) = \overrightarrow{\mathrm{grad}}\varphi(\vec{r})$$

le vecteur d'onde local. L'onde est **localement plane** si $\vec{k}(\vec{r})$ et $A(\vec{r})$ varient peu, en valeur relative, sur des distances de l'ordre de $\lambda = \dfrac{2\pi}{k}$ (domaine non grisé sur la figure 31).

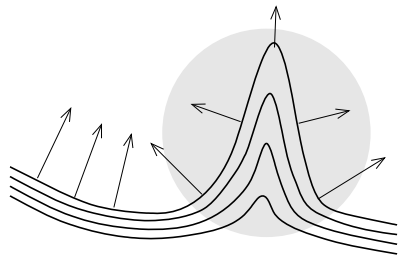


FIGURE 31

Alors $\vec{k}(\vec{r})$ vérifie la relation de dispersion locale $k(\vec{r}) = |\overrightarrow{\nabla}\varphi| = n(\vec{r})\,\dfrac{\omega}{c}$. Si on appelle **rayon** une courbe γ_0 tangente en tout point à $\vec{k}(\vec{r})$, $\vec{k}$ donnant la direction de propagation de l'énergie, on en déduit pour γ_0 à la fois le **principe de Fermat** et la **loi de Descartes** :

$$\int_A^B k(\vec{r})\,\mathrm{d}l \text{ ou } \int_A^B n(\vec{r})\,\mathrm{d}l \text{ **minimum local**} \quad ; \quad \frac{\mathrm{d}\vec{k}}{\mathrm{d}l} = \overrightarrow{\mathrm{grad}}\,k(\vec{r}) \text{ ou } \frac{\mathrm{d}}{\mathrm{d}l}\,n\hat{u} = \overrightarrow{\mathrm{grad}}\,n\,.$$

DÉMONSTRATION : pour Fermat (figure 32) on écrit

$$\int_{\gamma_0} |\vec{k}(\vec{r})|\,\mathrm{d}l = \int_{\gamma_0} \overrightarrow{\nabla}\varphi\cdot\overrightarrow{\mathrm{d}l} = \varphi(B)-\varphi(A) = \int_{\gamma} \overrightarrow{\nabla}\varphi\cdot\overrightarrow{\mathrm{d}l} < \int_{\gamma} |\overrightarrow{\nabla}\varphi|\,\mathrm{d}l$$

(car $\overrightarrow{\nabla}\varphi$ est parallèle à $\overrightarrow{\mathrm{d}l}$ sur γ_0 mais pas sur toute autre trajectoire γ allant de A à B) ; pour Descartes on écrit

$$\frac{\mathrm{d}k_x}{\mathrm{d}l} = \partial_x k_x\,\frac{\mathrm{d}x}{\mathrm{d}l} + \partial_y k_x\,\frac{\mathrm{d}y}{\mathrm{d}l} + \partial_z k_x\,\frac{\mathrm{d}z}{\mathrm{d}l} = \frac{1}{2k}\,\partial_x(k_x^2+k_y^2+k_z^2) = \partial_x k$$

en utilisant $\dfrac{\mathrm{d}x}{\mathrm{d}l} = \dfrac{k_x}{k}\cdots$ ($\vec{k}$ parallèle à $\overrightarrow{\mathrm{d}l}$ sur γ_0) et $\partial_y k_x = \partial_x k_y\cdots$ (égalité de Schwarz pour φ).

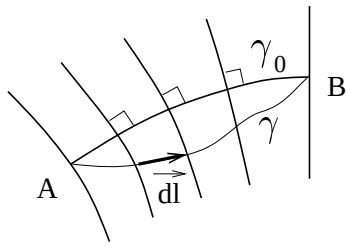


FIGURE 32

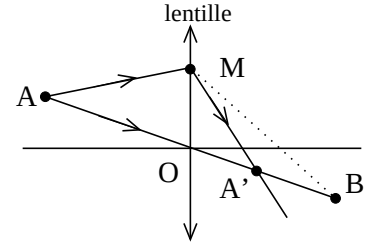


FIGURE 33

REMARQUES 1) On a supposé que $\overrightarrow{\nabla}\varphi$ est non singulier. Le rayon γ_0 étant donné, ceci peut toujours être satisfait localement mais pas globalement ; si tous les rayons issus de A convergent en un point A' situé avant B sur γ_0, le chemin optique de A à B n'est alors plus minimum mais simplement stationnaire ; sur la figure 33 le chemin optique (AMB) est plus court que $(AOA'B) = (AMA') + (A'B)$.
2) Si $\psi(t,\vec{r}) = A(t,\vec{r})\exp i\varphi(t,\vec{r})$ est localement une onde plane, on pose $\omega = -\partial_t\varphi$ et $\vec{k} = \overrightarrow{\nabla}\varphi$. Soit $\omega = f(\vec{k},t,\vec{r})$ la relation de dispersion (milieu linéaire mais non isotrope, non stationnaire et non homogène), soit $\vec{v}_g = \partial_{\vec{k}}\omega$ la vitesse de groupe et soit $\mathrm{d}_t = \partial_t + \vec{v}_g\,.\,\overrightarrow{\nabla}$ une "dérivée particulaire" ($\mathrm{d}_t\vec{r} = \vec{v}_g$). Un bon exercice de dérivées partielles consiste à montrer que $\mathrm{d}_t\omega = \partial_t f$ et $\mathrm{d}_t\vec{k} = -\overrightarrow{\nabla} f$ (relations qui généralisent les lois de Descartes et relient **optique géométrique et mécanique** ; *cf.* section 9.1).

▷ Caustiques et problèmes d'extremum

Soient dans l'air ($n = 1$) une surface d'onde $z = f(x,y)$ (point courant $m(x,y,z)$) et $M(X,Y,Z)$ un point fixe extérieur. On cherche les rayons m^*M passant par M, c.a.d. les droites qui rendent la distance mM extrémale ($\mathrm{d}(mM) = -\hat{u}_m\cdot\overrightarrow{\mathrm{d}m} = 0$ pour m^*).

Les caustiques (enveloppes de rayons) correspondent aussi aux lieux géométriques des points M où le problème d'extremum admet une **bifurcation** (c.a.d. un changement du nombre de solutions pour m^*).

EXEMPLE 1. $f(x,y) = \dfrac{x^2}{2R} + A\,\dfrac{x^3}{6}$ ($A > 0$; figure 30). Pour M près de la normale en O à la surface (axe Oz) on a $mM = \sqrt{(Z-f)^2 + (X-x)^2 + (Y-y)^2} \simeq Z - f + \dfrac{1}{2Z}\big((X-x)^2 + (Y-y)^2\big)$. x^* et y^* vérifient $-\dfrac{A}{2}\,x^{*2} + x^*(Z^{-1} - R^{-1}) - Z^{-1}X = 0$ et $y^* = Y$, d'où à priori 0 ou 2 solutions. Le cas limite de la racine double $x_1^* = x_2^*$ définit la caustique (paraboloïde) d'équation $(Z-R)^2 - 2AR^3X = 0$. Si $A = 0$, la caustique est la droite $Z = R$, $X = 0$ (il y a alors une infinité de solutions pour x^*).

EXEMPLE 2 (figure 34a). $f(x,y) = \frac{x^2}{2R_1} + \frac{y^2}{2R_2}$ ($R_1 \simeq R_2 > 0$). Pour M près de Oz, x^* et y^* sont donnés par $x^*(Z^{-1} - R_1^{-1}) = Z^{-1}X$ et $y^*(Z^{-1} - R_2^{-1}) = Z^{-1}Y$; les caustiques sont les deux droites de directions perpendiculaires Δ_1 ($Z = R_1$, $X = 0$) et Δ_2 ($Z = R_2$, $Y = 0$) ; elles se réduisent au foyer de l'onde $(0,0,R)$ si $R_1 = R_2 = R$. En pratique la surface d'onde est limitée par une pupille circulaire $x^2 + y^2 = r_{\max}^2$. Le calcul précédent montre que les rayons qui en sont issus coupent alors un plan $Z = \mathrm{C^{ste}}$ à l'intérieur de l'ellipse $\frac{X^2R_1^2}{(R_1-Z)^2} + \frac{Y^2R_2^2}{(R_2-Z)^2} = r_{\max}^2$. Cette ellipse est un cercle de rayon $r_{\max}\frac{|R_2-R_1|}{2R_0}$ dans le plan de convergence "moyen" $Z = \frac{R_1+R_2}{2} = R_0$, et se réduit à des segments de longueur $\frac{2r_{\max}|R_2-R_1|}{R_0}$ sur les caustiques (figure 34b).

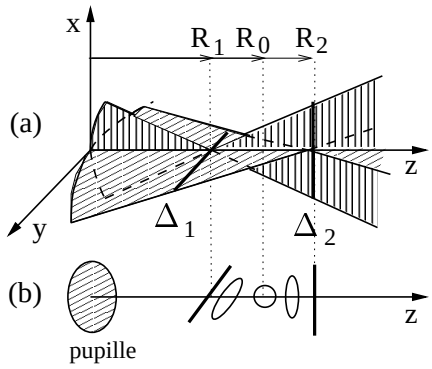


FIGURE 34

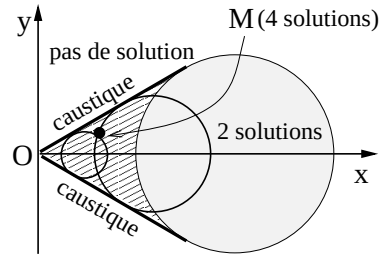


FIGURE 35

EXEMPLE 3. $f(x,y) = -Bx_0x(x^2+y^2) + \frac{x^2+y^2}{2d}$ ($B > 0$). Cherchons la trace de la caustique sur le plan $Z = d$ (plan de convergence de l'onde sphérique si $B = 0$) près de l'axe Oz. m^* s'obtient en extrémalisant $Bx_0x(x^2+y^2) - \frac{xX+yY}{d}$, ce qui donne $\frac{X}{d} = Bx_0r^{*2}(2+\cos 2\varphi^*)$ et $\frac{Y}{d} = Bx_0r^{*2}\sin 2\varphi^*$ en posant $x^* = r^*\cos\varphi^*$ et $y^* = r^*\sin\varphi^*$. Dans le plan $Z = d$, les cercles $r^* = \mathrm{C^{ste}}$ centrés en $(X = 2Bx_0r^{*2}, Y = 0)$ et de rayon Bx_0r^{*2} remplissent la région $X > \sqrt{3|Y|}$ (figure 35). Dans cette région il y a pour tout point M quatre solutions m^* (deux valeurs de r^* et pour chacune les solutions φ^* et $\varphi^* + \pi$). A l'extérieur de cette région ($X < \sqrt{3|Y|}$) il n'y a pas de solution. La caustique est donc formée des deux demi-droites $X = \sqrt{3|Y|}$ faisant entre elles un angle de 60^o. (En pratique une pupille $r = r_{max}$ limite la caustique et introduit une "sous région" où il y a deux solutions.)

▷ Aberrations géométriques d'un système optique centré

La figure 36 rappelle la situation expérimentale (*cf.* cours d'optique pour plus de détails).

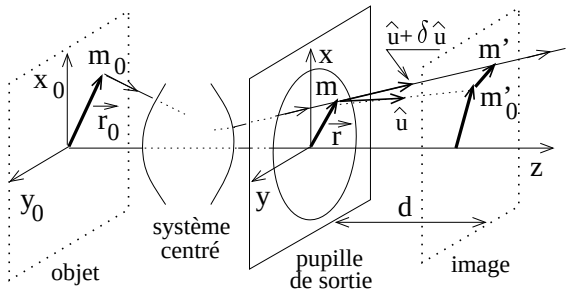


FIGURE 36

Si m'_0 était l'image rigoureuse de m_0, le chemin optique de m_0 à un point m de la pupille de sortie s'écrirait $L(m_0, m) = \mathrm{C}^{\mathrm{ste}} - mm'_0$, et le vecteur unitaire $\hat{u}$ du rayon en m serait parallèle à $\overrightarrow{mm'_0}$. Les corrections à L s'écrivent à priori

$$\delta L = \tfrac{A}{4}\,(\vec{r}\cdot\vec{r})^2 + B(\vec{r}\cdot\vec{r})\,(\vec{r}\cdot\vec{r}_0) + \tfrac{1}{2}\big(C(\vec{r}\cdot\vec{r}_0)^2 + C'(\vec{r}\cdot\vec{r})\,(\vec{r}_0\cdot\vec{r}_0)\big) + D(\vec{r}\cdot\vec{r}_0)\,(\vec{r}_0\cdot\vec{r}_0)$$

en vertu de la symétrie de révolution du système ; (il n'y a pas de scalaires du troisième degré et ceux du second degré ont été déjà pris en compte dans l'approximation de Gauss). On en déduit la correction d'orientation du rayon sortant :

$$\delta\hat{u} \simeq \delta\hat{u}_\perp = \partial_{\vec{r}}\,\delta L = Ar^2\vec{r} + B\left(2(\vec{r}\cdot\vec{r}_0)\,\vec{r} + r^2\vec{r}_0\right) + \big(C(\vec{r}\cdot\vec{r}_0)\vec{r}_0 + C' r_0^2\vec{r}\big) + D r_0^2\vec{r}_0\ .$$

Ce rayon coupe le plan image en m' tel que $\overrightarrow{m'_0m'} \simeq \delta\hat{u}\,d$ (aberration). A définit l'**aberration de sphéricité** (courbure de la surface d'onde plus grande en A_2 qu'en A_1 sur la figure 37a) ; B définit celle de **coma** (analogue de l'exemple 3 ci-dessus avec $\vec{r}_0 \equiv (x_0, 0)$), C (analogue de $\frac{1}{R_1} - \frac{1}{R_2}$ dans l'exemple 2) celle d'**astigmatisme**, $C' + \frac{C}{2}$ celle de **courbure de champ** (responsable de l'écart entre le point de convergence moyen et le plan image), et enfin D celle de **distorsion** (correction $D\,r_0^2\,d$ au grandissement ; la figure 37b représente l'image "distordue" d'un carré).

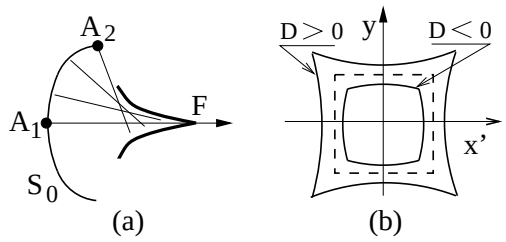


FIGURE 37

7.4 APPLICATIONS À LA THERMODYNAMIQUE

Changements de variables, définition de grandeurs importantes par des dérivées partielles, égalité de Schwarz, relation d'Euler, variations secondes, méthode des multiplicateurs de Lagrange, *etc.* figurent parmi les outils fréquemment utilisés en thermodynamique. Un choix d'exemples qui reprend quelques grandes problématiques de cette discipline est brièvement présenté (*cf.* aussi les sections 1.1.4, 5.1.4 et 10.2.3).

7.4.1 Rôle clé de l'entropie ; équations d'état ; coexistence de phases

▷ Coefficients élastiques et thermiques

La thermodynamique a commencé avec l'étude expérimentale des équations d'état de fluides $f(P,V,T)=0$ ($PV=RT$ pour une mole de gaz parfait (G.P.)) et celle des apports thermiques réversibles. Les coefficients élastiques α , β et χ_T sont définis par

$$\alpha=\frac{1}{V}\,\frac{\partial V}{\partial T}\Big|_P \quad , \quad \beta=\frac{1}{P}\,\frac{\partial P}{\partial T}\Big|_V \quad , \quad \chi_T=-\frac{1}{V}\,\frac{\partial V}{\partial P}\Big|_T \ ,$$

c'est-à-dire $\dfrac{\mathrm{d}V}{V}=\alpha\,\mathrm{d}T-\chi_T\,\mathrm{d}P$ et $\mathrm{d}P=\beta\,P\,\mathrm{d}T-\dfrac{1}{\chi_T}\,\dfrac{\mathrm{d}V}{V}$, et les coefficients thermiques c_V, l, c_P et h ainsi que χ_S sont définis par ($\delta Q_{\rm rev}$ forme différentielle) :

$$\delta Q_{\rm rev}=c_V\,\mathrm{d}T+l\,\mathrm{d}V=c_P\,\mathrm{d}T+h\,\mathrm{d}P \quad ; \quad \chi_S=-\frac{1}{V}\,\frac{\mathrm{d}V}{\mathrm{d}P}\ \text{à}\ \delta Q_{\rm rev}=0\ .$$

Tous ces coefficients ne sont pas indépendants. En écrivant $\mathrm{d}f=\partial_V f\,\mathrm{d}V+\partial_P f\,\mathrm{d}P+\partial_T f\,\mathrm{d}T=0$, on exprime α , β et χ_T en fonction des trois dérivées partielles de f (par exemple $\alpha V=-(\partial_T f)\,(\partial_V f)^{-1}$), et on obtient $\alpha=P\,\beta\,\chi_T$. De même en portant $\mathrm{d}V$ (resp. $\mathrm{d}P$) dans la première (resp. la seconde) expression de $\delta Q_{\rm rev}$, il vient $c_P-c_V=l\,\alpha\,V=-h\,\beta\,P$. Enfin $\delta Q_{\rm rev}=0$ entraîne $\dfrac{\mathrm{d}V}{\mathrm{d}P}=\dfrac{c_V}{l}\,\dfrac{h}{c_P}$ d'où $\dfrac{\chi_T}{\chi_S}=\dfrac{c_P}{c_V}$. A ce niveau de description les trois fonctions $P(T,V)$, $c_V(T,V)$ et $l(T,V)$ sont indépendantes.

▷ Equations d'état et entropie : gaz parfait ; gaz réel ; corps noir

La différentielle de $S(U,V)$ (U énergie interne, V volume)

$$\mathrm{d}S=\frac{1}{T}\,(\mathrm{d}U+P\,\mathrm{d}V)$$

(équivalent à $\mathrm{d}U=-P\,\mathrm{d}V+\delta Q_{\rm rev}$ avec $\delta Q_{\rm rev}=T\,\mathrm{d}S$) définit la **température thermodynamique** $T(U,V)=\dfrac{\partial S}{\partial U}\Big|_V^{-1}$ et la pression $P(U,V)=T\,\dfrac{\partial S}{\partial V}\Big|_U$. On déduit donc de S à la fois l'équation d'état $P(T,V)$, l'énergie interne $U(T,V)$ et les coefficients thermiques. Trois exemples explicites sont :

$$S=R\,\ln V+f(U) \quad ; \quad S=R\,\ln(V-b)+f\left(U+\frac{a}{V}\right) \quad ; \quad S=\mathrm{C}^{\rm ste}\,U^{\frac{3}{4}}V^{\frac{1}{4}}\ .$$

Pour le premier (une mole de G.P.), $R\,\ln V$ mesure la liberté de position des molécules et $f(U)$ celle associée à la répartition de l'énergie U entre elles. Dans le second (une mole de gaz réel), on tient compte de la taille des molécules (restriction du volume "libre") et de l'énergie potentielle (liée aux interactions à distance) $-\dfrac{a}{V}$, négative et proportionnelle à la densité. La troisième expression (corps noir, c.a.d. gaz de photons) a été introduite aux sections 1.1.4. et 1.3.2. Pour le G.P. la relation $\mathrm{d}S=RV^{-1}\,\mathrm{d}V+f'(U)\,\mathrm{d}U$ donne $PV=RT$ (donc T coïncide bien avec la température du gaz parfait) et $f'(U)=T^{-1}$ (relation qui définit $U_{\rm GP}(T)$) ; on a $\alpha=\beta=T^{-1}$, $\chi_T=P^{-1}$, $l=P$, $c_P-c_V=R$ et $c_V=-T^2f''(U)$ (fonction de T). Pour le gaz réel un calcul aussi simple donne $\left(P+\dfrac{a}{V^2}\right)(V-b)=RT$ (**équation de Van der Waals**) et $U=U_{\rm GP}(T)-\dfrac{a}{V}$; enfin

pour le corps noir $\dfrac{\mathrm{d}S}{S} = \dfrac{3}{4}\dfrac{\mathrm{d}U}{U} + \dfrac{1}{4}\dfrac{\mathrm{d}V}{V}$ donne $T = \dfrac{4U}{3S}$ et, pour la pression de radiation $P = \dfrac{U}{3V} \propto T^4$ (**loi de Stefan**).

REMARQUE. La connaissance de $P(T,V)$ entraîne celles de $l(T,V)$ et de $\partial_V c_{_V}\big|_{_T}$; pour obtenir $c_{_V}(T,V)$ et donc S, il reste à déterminer la dépendance en T de $c_{_V}$. Démonstration : appliquée à l'énergie libre $F = U - TS$ et à $S(T,V)$ dont les différentielles s'écrivent

$$\mathrm{d}F(T,V) = -S\,\mathrm{d}T - P\,\mathrm{d}V \quad \text{et} \quad \mathrm{d}S(T,V) = T^{-1}(c_{_V}\,\mathrm{d}T + l\,\mathrm{d}V)\ ,$$

l'égalité de Schwarz donne $l = T\,\partial_T P\big|_{_V}$ et $\partial_V c_{_V}\big|_{_T} = T\,\partial_T(T^{-1}l)\big|_{_V}$. En considérant la différentielle $\mathrm{d}G(T,P) = V\,\mathrm{d}P - S\,\mathrm{d}T$ de l'**enthalpie libre** $G = F + PV$ et $\mathrm{d}S(T,P)$, on établit des relations analogues pour $h(T,P)$ et $c_{_P}(T,P)$.

▷ Systèmes ouverts ; extensivité de l'entropie ; enthalpie libre

Pour une phase contenant n_i moles des espèces chimiques i, l'entropie dépend des grandeurs extensives $\underline{X} \equiv (U,V,n_i)$. L'expression $\mathrm{d}S = T^{-1}\,(\mathrm{d}U + P\,\mathrm{d}V - \sum_i \mu_i\,\mathrm{d}n_i)$ définit les **potentiels chimiques** d'équilibre μ_i. L'extensivité $S(k\underline{X}) = k\,S(\underline{X})$ de S entraîne que S est une fonction homogène de degré un et que (relation d'Euler)

$$S = \sum_j X_j\,\partial_j S = T^{-1}\Big(U + PV - \sum_i \mu_i n_i\Big)\ ,$$

soit, à l'aide de l'enthalpie libre :

$$G = U + PV - TS = \sum_i \mu_i n_i\ .$$

Elle entraîne aussi que $S(\underline{X}) = n\,S(\underline{x})$ avec $n = \sum_i n_i$ et $\underline{x} = n^{-1}\underline{X}$ (**grandeurs molaires**). Les dérivées partielles de $S(\underline{X})$ satisfont $f(k\underline{X}) = f(\underline{X})$; ces fonctions homogènes de degré zéro, qui ne dépendent que $\underline{x}$, sont des **grandeurs intensives**.

EXEMPLES. Pour une seule espèce chimique :

$$S(U,V,n) = n\,S\Big(\frac{U}{n},\frac{V}{n},1\Big) = n\,s(u,v) \quad \text{et} \quad G(P,T,n) = n\,\mu(P,T)\ .$$

Pour le corps noir, S est bien extensive ($S(kU,kV) = k\,S(U,V)$), mais comme n n'apparaît pas, on a $\mu = 0$ et $G = 0$.

REMARQUE. Si au lieu de U, V et n_i on choisit les variables P, T et n_i, toutes les grandeurs extensives (U, V, S, G...) satisfont $f(P,T,kn_i) = k\,f(P,T,n_i)$, d'où $f = \sum_i n_i\,f_i$ avec $f_i = \partial_{n_i} f|_{P,T}$ (relation d'Euler pour les variables n_i). On a un **mélange parfait** si chaque f_i ne dépend que de P, T et n_i et pas des $n_{j\neq i}$.

▷ Concavité de l'entropie ; stabilité thermodynamique

Soient deux parties d'un même fluide initialement séparées et ayant chacune n_i ($i = 1,2$) moles et des grandeurs molaires $\underline{x}_i \equiv (u_i, v_i)$. Si $n_1\underline{x}_1 + n_2\underline{x}_2$ étant fixé (système isolé) l'évolution des deux parties se fait vers un même état d'équilibre (cas d'une seule phase), le second principe entraîne que :

$$(n_1+n_2)\,s\left(\frac{n_1\underline{x}_1 + n_2\underline{x}_2}{n_1+n_2}\right) \geq n_1\,s(\underline{x}_1) + n_2\,s(\underline{x}_2)\ .$$

Ceci implique la concavité de s (donc aussi celle de S pour n fixé) : le segment joignant deux points de coordonnées (u_i, v_i, s_i) de la surface $s(\underline{x})$ est situé en dessous d'elle. On en déduit (*cf.* section 7.1.1) la condition (dite de **stabilité thermodynamique**) :

$$\delta^2 S(U,V) = \frac{1}{2}\left(\mathrm{d}\Big(\frac{1}{T}\Big)\,\mathrm{d}U + \mathrm{d}\Big(\frac{P}{T}\Big)\,\mathrm{d}V\right) = \frac{1}{2T}\left(\mathrm{d}P\,\mathrm{d}V - \mathrm{d}T\,\mathrm{d}S\right) < 0\ .$$

Conséquences : en annulant successivement $\mathrm{d}T$, $\mathrm{d}S$, $\mathrm{d}V$ et $\mathrm{d}P$ dans l'inégalité on obtient que $\chi_{_T}$, $\chi_{_S}$, $c_{_V}$ et $c_{_P}$ sont positifs. Comme $\mathrm{d}G = V\,\mathrm{d}P - S\,\mathrm{d}T$ on a aussi :

$$\delta^2 G(T,P) = \frac{1}{2}\left(\mathrm{d}P\,\mathrm{d}V - \mathrm{d}T\,\mathrm{d}S\right) < 0\ .$$

EXEMPLE : **fluide de Van der Waals**. La considération des termes du second ordre en $\mathrm{d}U$ et $\mathrm{d}V$ dans $S(U+\mathrm{d}U, V+\mathrm{d}V) = R\,\ln(V+\mathrm{d}V-b) + f\left(U+\mathrm{d}U+a(V+\mathrm{d}V)^{-1}\right)$ donne :

$$\delta^2 S = \Big(-\tfrac{1}{2}\,\tfrac{R}{(V-b)^2} + \tfrac{a}{TV^3}\Big)\ \mathrm{d}V^2 + \tfrac{1}{2}\,f''(\varphi)\,\mathrm{d}\varphi^2 \quad (\varphi = U + \tfrac{a}{V} = U_{\mathrm{GP}}(T))\ .$$

Le terme en $\mathrm{d}\varphi^2$ est négatif (en admettant la concavité de S pour le G.P.). Celui en $\mathrm{d}V^2$ s'écrit $\frac{1}{2}\,\partial_V P|_T\,\mathrm{d}V^2$ (*cf.* équation d'état) et a le signe de $-\chi_{_T}$. Dans le plan (V,P) la région $\delta^2 S < 0$ est celle où les isothermes ($\mathrm{d}\varphi = 0$) ont une pente négative (région physiquement observable, y compris les zones de retard à la vaporisation et à la liquéfaction ; *cf.* figure 2a du Chapitre 5) ; c'est aussi la région non grisée sur la surface $S(U,V)$ (figure 38).

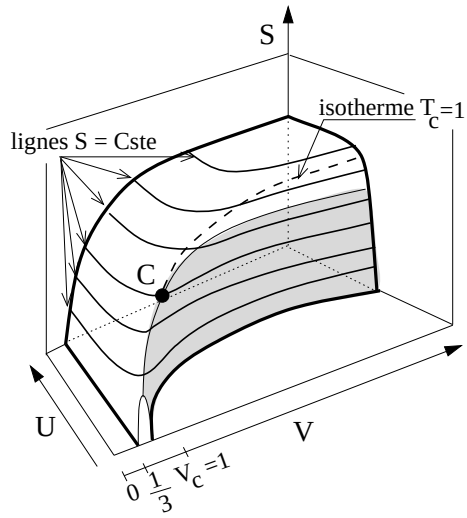


FIGURE 38

▷ Coexistence de phases

Avec les conditions $\mathrm{d}(n_1+n_2) = 0$ et $\mathrm{d}(n_1\underline{x}_1 + n_2\underline{x}_2) = 0$ ($\underline{x} \equiv (u,v)$), l'état d'équilibre correspond au maximum de $n_1 s(\underline{x}_1) + n_2 s(\underline{x}_2)$ et s'obtient avec la méthode des multiplicateurs de Lagrange. En écrivant $\mathrm{d}\left(\sum_i n_i\ (s(u_i,v_i) + \lambda_n + \lambda_{_u} u_i + \lambda_{_v} v_i)\right) = 0$, on obtient les six relations (annulation des coefficients de $\mathrm{d}u_i$, $\mathrm{d}v_i$ et $\mathrm{d}n_i$)

$$\lambda_{_u} = -T_i^{-1}\ \ ,\ \ \lambda_{_v} = -P_i T_i^{-1}\ \ \text{et}\ \ s(u_i,v_i) + \lambda_n + \lambda_{_u} u_i + \lambda_{_v} v_i = 0\ ,$$

d'où $T_1 = T_2$, $P_1 = P_2$ et $g_1 = g_2$ (ou $\mu_1 = \mu_2$). Ces trois équations admettent toujours la solution $\underline{x}_1 = \underline{x}_2$ mais, suivant le mélange initial, elles peuvent avoir une autre solution $\underline{x}_1 \neq \underline{x}_2$ (coexistence de deux phases) d'entropie plus grande. C'est elle qui conduit pour le fluide de Van der Waals à la **construction de Maxwell** des paliers d'équilibre gaz-liquide dans le plan (V, P). Cette construction revient (réflexion laissée au lecteur) à remplacer la surface de la figure 38 par son **enveloppe concave** (généralisation de la figure 6 du chapitre 5).

7.4.2 Potentiels thermodynamiques et équilibres

▷ Contraintes intérieures ; équilibre global ; variables associées

Un système contraint intérieurement peut être composé de plusieurs parties chacune en équilibre sans qu'il y ait équilibre global entre elles, par exemple plusieurs phases avec des parois adaptées permettant le contrôle des échanges volumiques, thermiques et moléculaires. Il peut aussi n'y avoir qu'un équilibre partiel au sein des phases (contrôle des réactions chimiques une par une dans chaque phase avec des catalyseurs adaptés). On peut aussi envisager des contraintes par la pensée si elles ne sont pas contradictoires avec les lois physiques. Dans tous les cas il faut distinguer dans l'expression de l'entropie totale S_t du système les variables $\underline{X}_t$ qui ne varient pas lors de l'évolution vers l'équilibre global de celles $\underline{X}$ dont les valeurs changent avec la levée des contraintes. $\underline{X}_t$ étant fixé l'équilibre global est atteint lorsque S_t, fonction de $\underline{X}$, est maximum.

EXEMPLE 1. **Equilibres chimiques**. (Système comportant c espèces chimiques i, φ phases α et siège de r réactions ρ). Pour chaque phase $\mathrm{d}S_\alpha = T_\alpha^{-1}\big(\mathrm{d}U_\alpha + P_\alpha\,\mathrm{d}V_\alpha - \sum_i \mu_{i\alpha}\,\mathrm{d}n_{i\alpha}\big)$. Si le système est isolé, $U_t = \sum_\alpha U_\alpha$ et $V_t = \sum_\alpha V_\alpha$ sont fixés. Si il est fermé, on peut séparer dans $\mathrm{d}n_{i\alpha}$ l'effet des réactions et celui des échanges entre phases : $\mathrm{d}n_{i\alpha} = \sum_\rho \nu_{i\rho}\,\mathrm{d}\xi_{\rho\alpha} + \delta n_{i\alpha}$; $\xi_{\rho\alpha}$ est l'**avancement de la réaction** ρ dans la phase α, $\nu_{i\rho}$ est le coefficient stoechiométrique de l'espèce i dans la réaction ρ, et on a les conditions $\sum_\alpha \delta n_{i\alpha} = 0$. L'équilibre global, $S_t = \sum_\alpha S_\alpha$ maximum, s'obtient en introduisant les $2 + c$ multiplicateurs de Lagrange $\lambda_U = -T^{-1}$, $\lambda_V = -T^{-1}P$ et $\lambda_{n_i} = T^{-1}\mu_i$. De

$$\mathrm{d}S_t - \frac{1}{T}\sum_\alpha \mathrm{d}U_\alpha - \frac{P}{T}\sum_\alpha \mathrm{d}V_\alpha + \sum_i \frac{\mu_i}{T}\,\mathrm{d}\left[\sum_\alpha\Big(n_{i\alpha} - \sum_\rho \nu_{i\rho}\xi_{\rho\alpha}\Big)\right] = 0\ ,$$

on déduit (annulation des coefficients de $\mathrm{d}U_\alpha$, $\mathrm{d}V_\alpha$, $\mathrm{d}n_{i\alpha}$ et $\mathrm{d}\xi_{\rho\alpha}$) :

$$T_\alpha = T \quad ; \quad P_\alpha = P \quad ; \quad \mu_{i\alpha} = \mu_i \quad ; \quad -\sum_i \mu_i \nu_{i\rho} = A_\rho = 0\ .$$

Ceci exprime l'égalité des températures, des pressions, des potentiels chimiques de chaque espèce et l'annulation des affinités A_ρ. Conséquence : si on s'intéresse aux $2 + \varphi(c-1)$ variables intensives P, T, et $n_{i\alpha}/n_\alpha$ décrivant l'ensemble des situations d'équilibre expérimentalement observables, les $(\varphi-1)c+r$ conditions sur les potentiels font que seules $v = c - r + 2 - \varphi$ d'entre elles sont indépendantes (**règle des phases** pour la variance).

EXEMPLE 2. **Loi de Boltzmann**. La répartition de N constituants microscopiques identiques et indépendants (pas nécessairement monoparticulaires) à raison de N_α par état microscopique α

(d'énergies ϵ_α fixées) peut se faire de $\Omega = \frac{N!}{\Pi_\alpha N_\alpha!}$ façons. On en déduit que $S(\underline{X} \equiv \{N_\alpha\}) \simeq -\sum_\alpha k_B N_\alpha \ln \frac{N_\alpha}{N}$ (formule de Stirling), et que $\mathrm{d}S = -k_B \sum_\alpha \ln \frac{N_\alpha}{N}\,\mathrm{d}N_\alpha$ (car $\sum_\alpha \mathrm{d}N_\alpha = 0$). La répartition d'équilibre avec $U = \sum_\alpha N_\alpha \epsilon_\alpha$ et $N = \sum_\alpha N_\alpha$ fixés s'obtient comme dans l'exemple 1 par : $\mathrm{d}S - T^{-1} \sum_\alpha (\epsilon_\alpha - \mu)\,\mathrm{d}N_\alpha = 0$. En annulant les coefficients des $\mathrm{d}N_\alpha$, il vient $\frac{N_\alpha}{N} = \exp -\frac{1}{k_B T}(\epsilon_\alpha - \mu)$.

▷ Potentiels thermodynamiques

Souvent on s'intéresse à des évolutions relatives à des systèmes en équilibre de température et de pression avec une **source** qui fixe P et T (l'atmosphère par exemple), ou à des déplacements d'équilibre lorsque P et (ou) T varient. Il faut alors considérer l'entropie

$$S_t = S + T^{-1}(U_s + PV_s) = S + S_s$$

de l'ensemble isolé constitué par le système (d'entropie S) et la source (indicée par s) :

$$S_t = T^{-1}(U_t + PV_t - G) \quad \text{avec } U_t = U + U_s,\ V_t = V + V_s \text{ et } G = U + PV - TS\ .$$

S_t dépend de $\underline{X}_t \equiv (U_t, V_t, P, T)$ et, à travers G, de variables de contraintes internes au système $\underline{X}$. Si on écrit $\mathrm{d}S = T^{-1}(\mathrm{d}U + P\,\mathrm{d}V) + \partial_{\underline{X}} S\big|_{U,V}\,\mathrm{d}\underline{X}$, on a :

$$\mathrm{d}G = V\,\mathrm{d}P - S\,\mathrm{d}T - \underline{A}\,\mathrm{d}\underline{X} \ \text{ avec } \ \underline{A} = -\partial_{\underline{X}} G\big|_{P,T} = T\,\partial_{\underline{X}} S\big|_{U,V}\ .$$

En tant que fonction de $\underline{X}$, $G(P, T, \underline{X})$ s'appelle potentiel thermodynamique car toute évolution $\Delta S_t \geq 0$ de l'ensemble isolé système plus source (U_t et V_t fixés) s'accompagne d'une diminution de G (jusqu'à un minimum). A l'équilibre $\underline{A} = 0$ et $G_{\mathrm{eq}}(P,T) = G\big(P, T, \underline{X}_{\mathrm{eq}}(P,T)\big)$.

REMARQUES. 1) Si l'ensemble est thermiquement (mais pas mécaniquement) isolé, et si $\Delta V_t = 0$, la condition $\Delta S_t \geq 0$ devient $-\Delta U_t \leq -\Delta G$. Elle conduit à la notion de **travail utile maximum** $-\Delta G$ pouvant être extrait de l'ensemble lorsque $\underline{X}$ varie (exemple : travail électrique fourni par une pile). 2) Si la source ne fixe que la température, G doit être remplacé par $F = U - TS$.

▷ Déplacements d'équilibres

Une variation de P et (ou) T induit une variation de $\underline{X}_{\mathrm{eq}}$. En se limitant à une seule composante pour $\underline{X}$, ce déplacement s'obtient en minimisant (par rapport à X)

$$G(P + \mathrm{d}P, T + \mathrm{d}T, X) = G(P,T,X) + V(P,T,X)\,\mathrm{d}P - S(P,T,X)\,\mathrm{d}T\ ,$$

dont le développement près de $X_{\mathrm{eq}} = X_{\mathrm{eq}}(P,T)$ est la fonction du second degré en X :

$$G(P,T,X_{\mathrm{eq}}) - \frac{1}{2}\,\partial_X A\big|_{P,T}\,(X - X_{\mathrm{eq}})^2 + \Big[\partial_X V\big|_{P,T}\,\mathrm{d}P - \partial_X S\big|_{P,T}\,\mathrm{d}T\Big](X - X_{\mathrm{eq}})\ .$$

($\partial_X A\big|_{P,T} < 0$ puisque $G(P,T,X)$ est minimum en X_{eq}.)

EXEMPLE 1. **Une réaction chimique dans une phase** (fermée). Dans ce cas $\underline{X} = \xi$ est l'avancement de la réaction ($\mathrm{d}n_i = \nu_i \mathrm{d}\xi$), et $A = -\sum_i \mu_i \nu_i$ (on rappelle que $\mathrm{d}S = T^{-1}(\mathrm{d}U + P\,\mathrm{d}V - \sum_i \mu_i\,\mathrm{d}n_i)$). Pour $\mathrm{d}T = 0$, $\mathrm{d}\xi_{\mathrm{eq}}$ a le signe de $(-\partial_X V)\,\mathrm{d}P$: **loi de Le Chatelier** ; (la figure 39 représente le déplacement d'équilibre pour $\mathrm{d}P > 0$ et $V(P,T,\xi)$ fonction croissante de ξ). Pour $\mathrm{d}P = 0$, $\mathrm{d}\xi_{\mathrm{eq}}$ a le signe de $\partial_X S\big|_{P,T}\,\mathrm{d}T$ (ou encore de $\partial_X H\big|_{P,T}\,\mathrm{d}T$ car $\mathrm{d}(H = U + PV) = T\,\mathrm{d}S$ si $\mathrm{d}P = 0$ et si $\xi = \xi_{\mathrm{eq}}$) : **loi de Van't Hoff**.

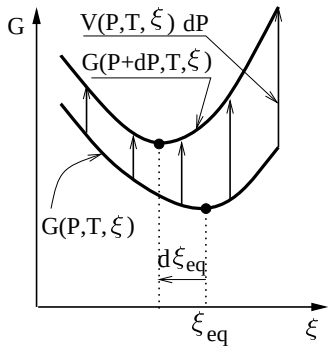


FIGURE 39

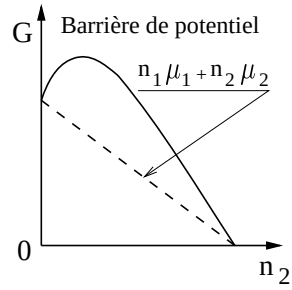


FIGURE 40

EXEMPLE 2. **Système diphasé** : $G = n_1\mu_1 + n_2\mu_2$ avec $n_1 = n - n_2$. Il y a équivalence avec une réaction $1 \rightleftharpoons 2$ ($\xi = n_2$ et $A = \mu_1 - \mu_2$), mais les phases étant séparées, μ_1 et μ_2 ne dépendent que de P et T et $\partial_\xi A|_{P,T} = 0$. Contrairement au cas U, V fixés (où $\partial_\xi A|_{U,V} \neq 0$), on a **évolution totale** vers la phase de plus petit μ (par exemple la glace, phase 2, si $\theta < 0^o\,C$). Remarque : l'observation, en apparence contradictoire, d'eau surfondue (phase 1) tient à ce que si on avait un petit glaçon dans l'eau, il faudrait tenir compte dans l'énergie totale U (et donc dans G) de l'énergie de tension superficielle $\gamma \times$ (surface du glaçon) $\propto (n_2)^{\frac{2}{3}}$. Alors, près de $n_2 = 0$, $G(P, T, n_2)$ croit au lieu de décroitre : il y a une **barrière de potentiel** (figure 40) qui explique la **métastabilité** de l'eau surfondue.

EXEMPLE 3 : **Bifurcation ; paramètre d'ordre**. Si le potentiel thermodynamique Φ (F ou G ou...) est tel que la variation d'un paramètre de controle $\underline{X}_t$ (P, T, champ magnétique $\overrightarrow{B}$, *etc.*) entraîne pour $\underline{X}$, au lieu d'un simple déplacement d'équilibre, une bifurcation de la solution d'équilibre $\underline{X}_{\text{eq}}$ (correspondant au minimum du potentiel), on dit que $\underline{X}$ est un paramètre d'ordre. La figure 41 donne, quand T varie, un exemple d'échange de deux solutions, et la figure 42 le passage de une à deux solutions "symétriques" l'une de l'autre. Comme $\frac{\mathrm{d}}{\mathrm{d}T} G\,(T, P, X_{\text{eq}}(T)) = \partial_T G = -S$ à P constant, le premier cas correspond à $S_2 \neq S_1$; un exemple explicite est fourni par le fluide de Van der Waals (avec T fixé et P variable) : $G(T, P, X \equiv V) = \big(-\frac{a}{V} + \varphi(T)\big) + PV - T\big(R\ln(V-b) + f(\varphi(T))\big)$. Le second cas correspond à $S_2 = S_1$; le ferromagnétisme traité à la section 5.1.4. (figure 10) en est un exemple (avec $\Phi \equiv F$ et $X \equiv x$ taux d'aimantation).

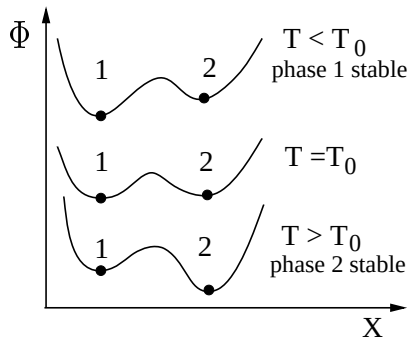


FIGURE 41

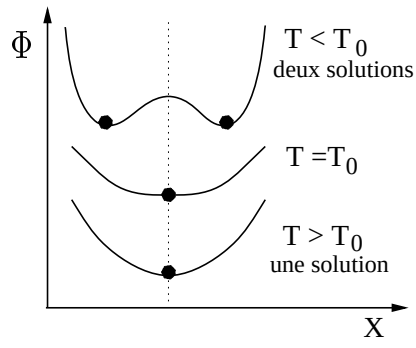


FIGURE 42

7.5 APPLICATIONS À L'ÉLECTROMAGNÉTISME

7.5.1 Formulation intégrale ; champs statiques ; milieux

▷ Electromagnétisme dans le vide

Rappelons les **équations de Maxwell** qui relient les champs électrique $\overrightarrow{E}$ et magnétique $\overrightarrow{B}$ aux densités volumiques de charge ρ et de courant $\vec{j}$:

$$\operatorname{div}\overrightarrow{E}=\frac{\rho}{\epsilon_0}\quad;\quad \operatorname{div}\overrightarrow{B}=0\quad;\quad \overrightarrow{\operatorname{rot}}\,\overrightarrow{E}=-\frac{\partial\overrightarrow{B}}{\partial t}\quad;\quad \overrightarrow{\operatorname{rot}}\,\overrightarrow{B}=\mu_0\left(\vec{j}+\epsilon_0\frac{\partial\overrightarrow{E}}{\partial t}\right).$$

Ces E.D.P. sont la version locale des relations intégrales (avec contours et surfaces fixes) :

$$\Phi_S(\overrightarrow{E})=\frac{Q}{\epsilon_0}\;;\;\Phi_S(\overrightarrow{B})=0\;;\;\mathcal{C}_\gamma(\overrightarrow{E})=-\frac{\mathrm{d}}{\mathrm{d}t}\varphi_s(\overrightarrow{B})\;;\;\mathcal{C}_\gamma(\overrightarrow{B})=\mu_0\left(I+\epsilon_0\frac{\mathrm{d}}{\mathrm{d}t}\varphi_s(\overrightarrow{E})\right).$$

Les notations φ_s, Φ_S et $\mathcal{C}_\gamma$ sont définies à la section 7.2.3. Q est la charge contenue dans le volume limité par la surface fermée S, et I est l'intensité du courant (flux $\varphi_s(\vec{j})$ ou quantité de charge par unité de temps) à travers la surface ouverte S qui s'appuie sur γ. Pour le circuit $\{\gamma-\gamma\}$ de la figure 43 (sur lequel s'appuie une surface fermée), la dernière relation s'écrit (puisque alors $\varphi\equiv\Phi$)

$$0=I+\frac{\mathrm{d}Q}{\mathrm{d}t}$$

(version intégrale de $\partial_t\rho+\operatorname{div}\vec{j}=0$) ; c'est cette **L.C. de la charge** que Maxwell a pris en compte pour introduire le **courant de déplacement** $\epsilon_0\,\partial_t\overrightarrow{E}$.

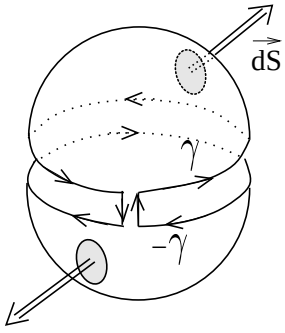


FIGURE 43

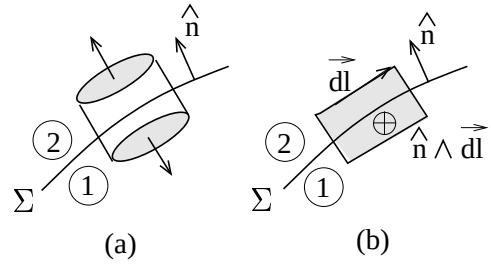


FIGURE 44

(Dis)continuités de $\overrightarrow{E}$ et $\overrightarrow{B}$. Soit Σ une surface porteuse de densités surfaciques de charge σ et de courant $\vec{j}_s$. Si V est un petit volume cylindrique (limité par S fermée) qui découpe sur Σ la surface élémentaire $\mathrm{d}\Sigma$ et dont la hauteur tend vers zéro (figure 44a), alors :

$$\Phi_S(\overrightarrow{E})=(E_{n2}-E_{n1})\,\mathrm{d}\Sigma=\frac{\sigma_s}{\epsilon_0}\,\mathrm{d}\Sigma\quad\text{et}\quad\Phi_S(\overrightarrow{B})=(B_{n2}-B_{n1})\,\mathrm{d}\Sigma=0\;;$$

on en déduit, pour les composantes normales des champs :

$$E_{n2} - E_{n1} = \frac{\sigma_s}{\epsilon_0} \quad ; \quad B_{n2} = B_{n1} \ .$$

De même si γ est un circuit élémentaire fermé dont les côtés qui coupent Σ tendent vers zéro, et les côtés parallèles à Σ sont $\overrightarrow{\mathrm{d}l}$ côté 2 et $-\overrightarrow{\mathrm{d}l}$ côté 1 (figure 44b), alors :

$$\mathcal{C}_\gamma(\overrightarrow{E}) = (\overrightarrow{E}_2 - \overrightarrow{E}_1)\cdot\overrightarrow{\mathrm{d}l} = 0 \ \text{ et } \ \mathcal{C}_\gamma(\overrightarrow{B}) = (\overrightarrow{B}_2 - \overrightarrow{B}_1)\cdot\overrightarrow{\mathrm{d}l} = \mu_0\, I = \mu_0\, \vec{j}_s\cdot(\hat{n}\wedge\overrightarrow{\mathrm{d}l}) \ .$$

Ceci entraîne pour les composantes tangentielles des champs :

$$\overrightarrow{E}_{t2} = \overrightarrow{E}_{t1} \quad ; \quad \overrightarrow{B}_{t2} - \overrightarrow{B}_{t1} = \mu_0\, \vec{j}_s \wedge \hat{n} \ .$$

La seule hypothèse faite est d'avoir supposé les champs bornés.

▷ Champs statiques ; approximation des régimes quasi-stationnaires (A.R.Q.S)

La formulation intégrale permet de déterminer ces champs très simplement en utilisant des "arguments de symétrie" (*cf.* section 3.4.3). On se limite à l'établissement des formules de base.

Electrostatique. Pour une charge ponctuelle q à l'origine, $\overrightarrow{E}(\vec{r})$ est dirigé selon $\hat{r}$. Donc pour une sphère S de rayon r centrée en O, on a $\Phi_S(\overrightarrow{E}) = 4\pi r^2 E = \dfrac{q}{\epsilon_0}$, d'où $\overrightarrow{E}(\vec{r}) = \dfrac{1}{4\pi\epsilon_0}\,\dfrac{q}{r^2}\,\hat{r}$.

Magnétostatique. On démontre facilement la **formule de Biot et Savart** donnant la contribution d'un élément de courant $I\overrightarrow{\mathrm{d}l}$ situé en m au champ magnétique en M

$$\overrightarrow{\mathrm{d}B} = \frac{\mu_0}{4\pi}\, I\, \frac{\overrightarrow{\mathrm{d}l}\wedge\hat{r}}{r^2} \quad \text{avec } \vec{r} = \overrightarrow{mM}$$

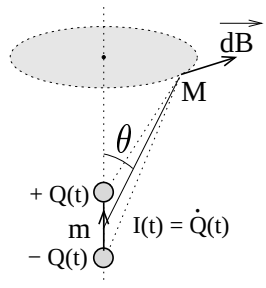


FIGURE 45

en remplaçant l'élément de courant par le dipole (ou l'antenne) de la figure 45. Comme $\overrightarrow{\mathrm{d}B}$ est orthoradial, on a $\mathcal{C}_\gamma(\overrightarrow{\mathrm{d}B}) = 2\pi r\,\sin\theta\,\mathrm{d}B = \epsilon_0\mu_0\,\dfrac{\mathrm{d}}{\mathrm{d}t}\varphi_s(\overrightarrow{\mathrm{d}E})$, avec dans l'approximation électrostatique $\varphi_s(\overrightarrow{\mathrm{d}E}) = \dfrac{1}{4\pi\epsilon_0}\,Q\big(2\pi(1-\cos\theta_+) - 2\pi(1-\cos\theta_-)\big) = \dfrac{1}{2\epsilon_0}\,Q\sin\theta\,\dfrac{\mathrm{d}l\,\sin\theta}{r}$,

d'où $\mathrm{d}B = \dfrac{\mu_0}{4\pi r^2}\, I \sin\theta\, \mathrm{d}l$. Un fil parcouru par un courant I peut être modélisé par une "suite continue" de tels dipoles. Remarque : si I donc $\overrightarrow{B}$ dépend du temps, il y a un champ électrique induit supplémentaire proportionnel à $\ddot{Q}$; il est négligeable si I varie lentement : A.R.Q.S. qui consiste à négliger les dérivées secondes par rapport au temps des grandeurs impliquées.

▷ Séparation macroscopique des charges et courants dans un milieu neutre (diélectrique et/ou magnétique) (figure 46)

On considère un milieu dans lequel les charges sont regroupées par "paquets neutres" (molécules par exemple) et, par la pensée, un élément de surface $\overrightarrow{\mathrm{d}S} = \mathrm{d}S\,\hat{n}$ à l'intérieur. Il va couper de nombreux "paquets" et laisser des deux côtés (1 et 2) de $\overrightarrow{\mathrm{d}S}$ des charges proportionnelles à $\mathrm{d}S$ et opposées, qu'on peut écrire $\pm\overrightarrow{P}\cdot\overrightarrow{\mathrm{d}S}$ ($\overrightarrow{P}\neq 0$ si le milieu est polarisé). Si $\overrightarrow{P}(\vec{r},t)$ est une fonction régulière de $\vec{r}$, on en déduit qu'un volume élémentaire d^3r dans le milieu contient une charge $-\mathrm{div}\,\overrightarrow{P}\,\mathrm{d}^3r$ (flux de $-\overrightarrow{P}$ à travers sa surface), et que l'intensité du courant associé au mouvement de ces charges à travers $\overrightarrow{\mathrm{d}S}$ est $\partial_t\overrightarrow{P}\cdot\overrightarrow{\mathrm{d}S}$.

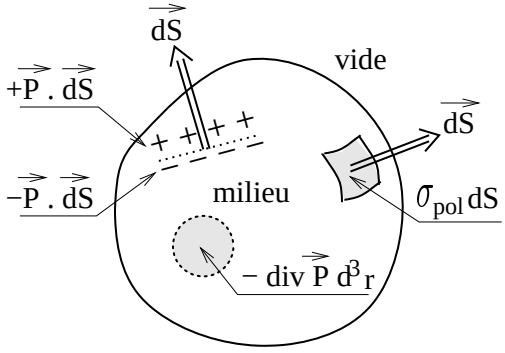


FIGURE 46

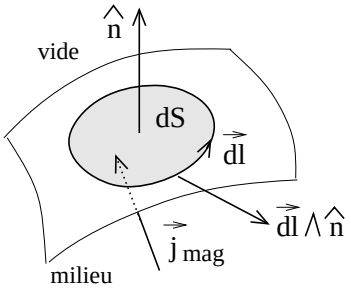


FIGURE 47

Le milieu peut donc être décrit par des **densités de charges et de courants** dites de **polarisation diélectrique** :

$$\sigma_{\mathrm{pol}} = \overrightarrow{P}\cdot\hat{n} \quad ; \quad \rho_{\mathrm{pol}} = -\mathrm{div}\,\overrightarrow{P} \quad ; \quad \vec{j}_{\mathrm{pol}} = \frac{\partial\overrightarrow{P}}{\partial t}\ .$$

($\sigma_{\mathrm{pol}}\,\mathrm{d}S$ est la charge laissée à la frontière du milieu quand $\overrightarrow{\mathrm{d}S}$ s'en rapproche.) La L.C. des charges $\partial_t\,\rho_{\mathrm{pol}} + \mathrm{div}\,\vec{j}_{\mathrm{pol}} = 0$ est satisfaite. Mais comme $\mathrm{div}\,\overrightarrow{\mathrm{rot}}\,(\cdots) = 0$, on peut cependant ajouter a priori à $\vec{j}_{\mathrm{pol}}$ une densité volumique $\overrightarrow{\mathrm{rot}}\,\overrightarrow{M}$ (qui devient surfacique à la frontière du milieu : $\overrightarrow{\mathrm{rot}}\,\overrightarrow{M}\cdot\overrightarrow{\mathrm{d}S} = \oint\overrightarrow{M}\cdot\overrightarrow{\mathrm{d}l} = \oint(\overrightarrow{M}\wedge\hat{n})\cdot(\overrightarrow{\mathrm{d}l}\wedge\hat{n})$; figure 47). Cette contribution, dite **magnétique** aux courants, est donc :

$$\vec{j}_{\mathrm{mag}} = \overrightarrow{\mathrm{rot}}\,\overrightarrow{M} \quad ; \quad \vec{j}_{\mathrm{s,mag}} = \overrightarrow{M}\wedge\hat{n}\ .$$

REMARQUE : en première approximation $\overrightarrow{P}$ est la densité volumique de moments dipolaires électriques ; par exemple (figure 48) pour une densité uniforme n de dipoles $\vec{p} = q\vec{l}$ identiques, on a $\overrightarrow{P} \cdot \overrightarrow{\mathrm{d}S} = nq\vec{l} \cdot \overrightarrow{\mathrm{d}S}$, car les dipoles coupés en deux par $\mathrm{d}S$ ont leurs centres situés dans un volume $\vec{l} \cdot \overrightarrow{\mathrm{d}S}$. De même $\overrightarrow{M}$ est assimilable à la densité volumique $n\vec{\mu}$ de moments magnétiques si on considère des "boucles microscopiques de courant" de moment $\vec{\mu} = i\vec{s}$ (figures 49a,b). En effet un élément $\mathrm{d}l$ traverse $n\,\overrightarrow{\mathrm{d}l} \cdot \vec{s}$ boucles auxquelles est associé le courant $ni\,\overrightarrow{\mathrm{d}l} \cdot \vec{s} = \overrightarrow{M} \cdot \overrightarrow{\mathrm{d}l}$, et donc tout circuit fermé γ est traversé par un courant $\oint_\gamma \overrightarrow{M} \cdot \overrightarrow{\mathrm{d}l} = \iint \overrightarrow{\mathrm{rot}}\, \overrightarrow{M} \cdot \overrightarrow{\mathrm{d}S}$.

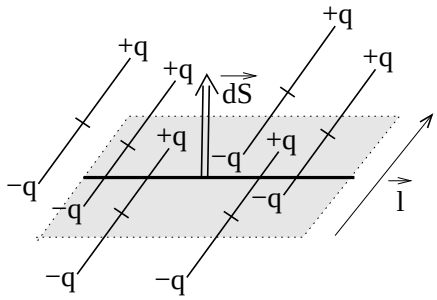


FIGURE 48

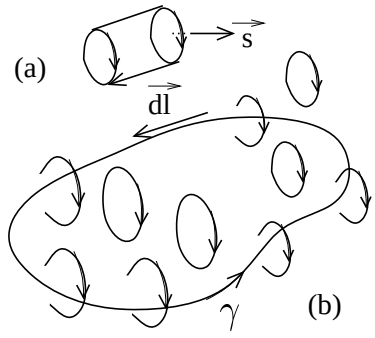


FIGURE 49

Equations de Maxwell. Les lois pour les champs macroscopiques $\overrightarrow{E}$ et $\overrightarrow{B}$ dans les milieux se déduisent alors de celles dans le vide en remplaçant Q par $Q_{\mathrm{ext}} - \Phi_S(\overrightarrow{P})$ et I par $I_{\mathrm{ext}} + \varphi_S\left(\frac{\partial \overrightarrow{P}}{\partial t} + \overrightarrow{\mathrm{rot}}\ \overrightarrow{M}\right)$, où Q_{ext} et I_{ext} sont les charges et courants extérieurs aux milieux. En posant $\overrightarrow{D} = \epsilon_0 \overrightarrow{E} + \overrightarrow{P}$ et $\overrightarrow{B} = \mu_0 (\overrightarrow{H} + \overrightarrow{M})$ elles s'écrivent simplement :

$$\Phi_S(\overrightarrow{D}) = Q_{\mathrm{ext}} \quad ; \quad \Phi_S(\overrightarrow{B}) = 0 \quad ; \quad \mathcal{C}_\gamma(\overrightarrow{E}) = -\frac{\mathrm{d}}{\mathrm{d}t}\varphi_S(\overrightarrow{B}) \quad ; \quad \mathcal{C}_\gamma(\overrightarrow{H}) = I_{\mathrm{ext}} + \frac{\mathrm{d}}{\mathrm{d}t}\varphi_S(\overrightarrow{D})\ .$$

Elles impliquent, à la traversée de toute surface S neutre séparant deux milieux, la continuité de D_n, B_n, $\overrightarrow{E}_t$ et $\overrightarrow{H}_t$ et, comme conséquence, celle du flux du vecteur de Poynting $(\overrightarrow{E} \wedge \overrightarrow{H}) \cdot \overrightarrow{\mathrm{d}S}$, qui correspond au **flux d'énergie électromagnétique**. Sous forme différentielle ces lois s'écrivent :

$$\mathrm{div}\, \overrightarrow{D} = \rho_{\mathrm{ext}} \quad ; \quad \mathrm{div}\, \overrightarrow{B} = 0 \quad ; \quad \overrightarrow{\mathrm{rot}}\ \overrightarrow{E} = -\partial_t\, \overrightarrow{B} \quad ; \quad \overrightarrow{\mathrm{rot}}\ \overrightarrow{H} = \vec{j}_{\mathrm{ext}} + \partial_t\, \overrightarrow{D}\ .$$

7.5.2 Potentiel scalaire et potentiel vecteur ; bilans d'énergie et de quantité de mouvement ; A.R.Q.S.

On considère les équations de Maxwell dans le vide.

▷ Potentiels scalaire et vecteur ; solution générale

Les conditions de nullité vues à la section 7.2.3 entraînent que $\mathrm{div}\, \overrightarrow{B} = 0$ a pour solution $\overrightarrow{B} = \overrightarrow{\mathrm{rot}}\ \overrightarrow{A}$, et que $\overrightarrow{\mathrm{rot}}\ \overrightarrow{E} = -\partial_t\, (\overrightarrow{\mathrm{rot}}\ \overrightarrow{A})$ permet d'écrire $\overrightarrow{E} = -\partial_t\, \overrightarrow{A} - \overrightarrow{\mathrm{grad}} V$. Les deux autres équations deviennent :

$$-\partial_t\, (\mathrm{div}\, \overrightarrow{A}) - \Delta V = \frac{\rho}{\epsilon_0}\ ,$$

$$\overrightarrow{\mathrm{rot}}\ (\overrightarrow{\mathrm{rot}}\ \overrightarrow{A}) = \overrightarrow{\mathrm{grad}}\, (\mathrm{div}\, \overrightarrow{A}) - \Delta \overrightarrow{A} = \mu_0 \big(\vec{j} + \epsilon_0 (-\partial_t^2 \overrightarrow{A} - \overrightarrow{\mathrm{grad}}\, \partial_t V)\big)\ .$$

Elles se simplifient grandement si on remarque que le **changement de jauge** sur les potentiels

$$\overrightarrow{A} \to \overrightarrow{A} + \overrightarrow{\text{grad}\varphi} \quad ; \quad V \to V - \partial_t \varphi$$

laisse les champs $\overrightarrow{E}$ et $\overrightarrow{B}$ inchangés quel que soit φ. En choisissant φ de sorte que le nouveau couple $(V, \overrightarrow{A})$ vérifie $\operatorname{div} \overrightarrow{A} + \epsilon_0\mu_0\, \partial_t V = 0$ (**jauge de Lorentz**), elles deviennent :

$$\left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) V = -\frac{\rho}{\epsilon_0} \quad ; \quad \left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) \overrightarrow{A} = -\mu_0 \vec{j} \qquad (\epsilon_0\mu_0 c^2 = 1)\ .$$

La solution "physique" de ces équations

$$V(\vec{r}_0, t_0) = \iiint \frac{1}{4\pi\epsilon_0} \frac{\rho\big(\vec{r}, t_0 - \frac{|\vec{r}-\vec{r}_0|}{c}\big)}{|\vec{r}-\vec{r}_0|}\, \mathrm{d}^3 r \quad ; \quad \overrightarrow{A}(\vec{r}_0, t_0) = \iiint \frac{\mu_0}{4\pi} \frac{\vec{j}\big(\vec{r}, t_0 - \frac{|\vec{r}-\vec{r}_0|}{c}\big)}{|\vec{r}-\vec{r}_0|}\, \mathrm{d}^3 r$$

est établie à la section 8.2.2. L'**A.R.Q.S.** ("∂_t^2 négligeable") équivaut à faire tendre c vers l'infini dans les E.D.P. ci-dessus et consiste donc à ne pas tenir compte de la propagation dans leur solution. Concrètement pour un phénomène de période T, l'A.R.Q.S. n'est valable que dans un domaine dont la taille L est telle que $L \ll cT$.

▷ Bilans d'énergie et de quantité de mouvement

Des équations de Maxwell dans le vide on déduit les relations

$$\iiint_V \vec{j}\cdot\overrightarrow{E}\, \mathrm{d}^3 r + \iiint_V \partial_t\Big(\epsilon_0 \frac{\overrightarrow{E}^2}{2} + \frac{\overrightarrow{B}^2}{2\mu_0}\Big)\, \mathrm{d}^3 r = -\oiint_S \Big(\overrightarrow{E} \wedge \frac{\overrightarrow{B}}{\mu_0}\Big)\cdot \overrightarrow{\mathrm{d}S}$$

$$\iiint_V \big(\rho\overrightarrow{E} + \vec{j}\wedge\overrightarrow{B}\big)\, \mathrm{d}^3 r + \iiint_V \epsilon_0 \partial_t(\overrightarrow{E}\wedge\overrightarrow{B})\, \mathrm{d}^3 r = -\oiint_S [\mathbf{P}]\, \overrightarrow{\mathrm{d}S}$$

avec $P_{xx} = \epsilon_0\Big(\dfrac{\overrightarrow{E}^2}{2} - E_x^2\Big) + \dfrac{1}{\mu_0}\Big(\dfrac{\overrightarrow{B}^2}{2} - B_x^2\Big)\cdots$ et $P_{xy} = -\epsilon_0 E_x E_y - \dfrac{1}{\mu_0} B_x B_y \cdots$.

DÉMONSTRATION : la première résulte de $\operatorname{div}(\overrightarrow{E}\wedge\overrightarrow{B}) = \overrightarrow{B}\,\overrightarrow{\text{rot}}\,\overrightarrow{E} - \overrightarrow{E}\,\overrightarrow{\text{rot}}\,\overrightarrow{B} = -\overrightarrow{B}\,\partial_t\overrightarrow{B} - \mu_0\overrightarrow{E}\,(\vec{j} + \epsilon_0\partial_t\overrightarrow{E})$. Pour la deuxième on déduit des E.D.P. de Maxwell que le membre de gauche n'est autre que l'intégrale triple de $[-\epsilon_0(\overrightarrow{E}\wedge\overrightarrow{\text{rot}}\,\overrightarrow{E} - \overrightarrow{E}\operatorname{div}\overrightarrow{E}) - \mu_0^{-1}(\overrightarrow{B}\wedge\overrightarrow{\text{rot}}\,\overrightarrow{B} - \overrightarrow{B}\operatorname{div}\overrightarrow{B})]$, car $\operatorname{div}\overrightarrow{B} = 0$; pour trouver $[\mathbf{P}]$ au second membre, il suffit d'écrire chaque composante de $[\cdots]$ comme une divergence ; par exemple : $(\overrightarrow{E}\wedge\overrightarrow{\text{rot}}\,\overrightarrow{E} - \overrightarrow{E}\operatorname{div}\overrightarrow{E})_x = E_y(\partial_x E_y - \partial_y E_x) - E_z(\partial_z E_x - \partial_x E_z) - E_x(\partial_x E_x + \partial_y E_y + \partial_z E_z) = \partial_x\frac{1}{2}(E_y^2 + E_z^2 - E_x^2) + \partial_y(-E_y E_x) + \partial_z(-E_z E_x)$. $[\mathbf{P}]$ est la **matrice de pression électromagnétique**.

Interprétation : comme pour une charge libre q soumise à l'action de $\overrightarrow{E}$ et $\overrightarrow{B}$ les variations de quantité de mouvement et d'énergie sont $\dfrac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q\,(\overrightarrow{E} + \vec{v}\wedge\overrightarrow{B})$ et $\dfrac{\mathrm{d}E_c}{\mathrm{d}t} = \vec{v}\cdot\dfrac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q\,\overrightarrow{E}\cdot\vec{v}$, il est naturel (en considérant le premier terme du membre de gauche) d'interpréter ces deux relations pour des charges libres comme :

"$\dfrac{\mathrm{d}}{\mathrm{d}t}$ [énergie ou quantité de mouvement, dans le volume V, des charges et du champ électromagnétique] $=-$ [flux sortant d'énergie ou de quantité de mouvement à travers la surface S qui limite V] ".

En particulier $\epsilon_0 \dfrac{\overrightarrow{E}^2}{2} + \dfrac{\overrightarrow{B}^2}{2\mu_0}$ est la densité volumique d'énergie électromagnétique du champ.

Approximation A.R.Q.S.. En écrivant $E^2 = \overrightarrow{E} \cdot (-\partial_t \overrightarrow{A} - \overrightarrow{\text{grad}} V)$ et $B^2 = \overrightarrow{B} \cdot \overrightarrow{\text{rot}}\, \overrightarrow{A}$, et en utilisant les équations de Maxwell, un calcul élémentaire conduit à :

$$\epsilon_0 \frac{E^2}{2} + \frac{B^2}{2\mu_0} = \frac{1}{2}\left(\rho V + \vec{j} \cdot \overrightarrow{A}\right) - \frac{1}{2}\,\text{div}(\epsilon_0 V \overrightarrow{E} + \mu_0^{-1} \overrightarrow{B} \wedge \overrightarrow{A}) + \frac{\epsilon_0}{2}\left(\overrightarrow{A} \cdot \partial_t \overrightarrow{E} - \overrightarrow{E} \cdot \partial_t \overrightarrow{A}\right) ,$$

$$\text{et} \quad -\vec{j} \cdot \overrightarrow{E} = \vec{j} \cdot \partial_t \overrightarrow{A} + V\, \partial_t \rho + \text{div}(V \vec{j}) \quad \text{(avec la L.C. de la charge)} .$$

Dans l'A.R.Q.S. les intégrales sur $\mathbb{R}^3$ des divergences donnent 0 (décroissance à l'infini au moins en r^{-2} et r^{-1} des champs et potentiels), et la dérivée temporelle de $\overrightarrow{A} \cdot \partial_t \overrightarrow{E} - \overrightarrow{E} \cdot \partial_t \overrightarrow{A}$ est négligeable. Le bilan énergétique devient :

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[E_{em} = \iiint \left(\frac{1}{2}\,\rho V + \frac{1}{2}\,\vec{j} \cdot \overrightarrow{A}\right) \mathrm{d}^3 r\right] = \iiint \left(V\, \partial_t \rho + \vec{j} \cdot \partial_t \overrightarrow{A}\right) \mathrm{d}^3 r .$$

On peut même séparer le **bilan électrique** du **bilan magnétique**.

E_{em} est facile à calculer pour des fils parcourus par des courants (magnétostatique). Comme pour chaque fil $\iiint_{\text{fil}} \vec{j} \cdot \overrightarrow{A}\, \mathrm{d}^3 r = I \oint_{\text{fil}} \overrightarrow{A} \cdot \overrightarrow{\mathrm{d}r} = I\, \varphi$ on obtient :

$$E_{em} = \textstyle\sum_i \frac{1}{2}\, I_i \varphi_i .$$

De même pour chaque fil $\iiint_{\text{fil}} \vec{j} \cdot \partial_t \overrightarrow{A}\, \mathrm{d}^3 r = I \oint_{\text{fil}} \partial_t \overrightarrow{A} \cdot \overrightarrow{\mathrm{d}r}$. Comme on a vu à la section 7.2.3 que $\oint_\gamma I \partial_t \overrightarrow{A} \cdot \overrightarrow{\mathrm{d}r} = I\, \frac{\mathrm{d}}{\mathrm{d}t} \varphi_S(\overrightarrow{B}) - \oint_\gamma I(\overrightarrow{\mathrm{d}r} \wedge \overrightarrow{B}) \cdot \vec{v}$, on en déduit le bilan général d'énergie

$$\mathrm{d}E_{em} = \textstyle\sum_i I_i\, \mathrm{d}\varphi_i - \delta W$$

où $-\delta W$, opposé du travail des forces de Laplace, est le travail des sources mécaniques et $\sum_i I_i\, \mathrm{d}\varphi_i$ est celui des sources électriques (*cf.* section 7.3.1). On vérifie que $-\delta W = \mathrm{d}E_p$ avec pour l'énergie potentielle $E_p = E_{em}$ ou $E_p = -E_{em}$ suivant que les paramètres choisis sont les flux ou les intensités.

7.5.3 Calculs avec des densités microscopiques ; rayonnement

On rappelle (*cf.* section 7.2.4) que pour des charges ponctuelles :

$$\rho^{\text{micro}}(\vec{r}, t) = \sum_i q_i\, \delta\big(\vec{r} - \vec{r}_i(t)\big) \quad ; \quad \vec{j}^{\,\text{micro}}(\vec{r}, t) = \sum_i q_i \vec{v}_i\, \delta\big(\vec{r} - \vec{r}_i(t)\big) .$$

Intégrées dans un volume V, ces densités donnent $\sum_i q_i$ et $\sum_i q_i \vec{v}_i$ où la somme porte sur les charges contenues dans V à l'instant t. Pour un volume élémentaire $\mathrm{d}^3 r$ centré en $\vec{r}$ contenant suffisamment de charges, $\sum_i q_i$ et $\sum_i q_i \vec{v}_i$ sont assimilables à $\rho(\vec{r}, t)\, \mathrm{d}^3 r$ et $\vec{j}(\vec{r}, t)\, \mathrm{d}^3 r$ (ou $I\, \overrightarrow{\mathrm{d}l}$ pour un élément de courant) avec ρ et $\vec{j}$ "continus". Nous allons voir quelques applications de ce "double langage" avant de considérer l'effet du lissage de ces densités (remplacement de $\delta(\vec{r})$ par $f(\vec{r})$).

▷ Moments dipolaire et quadrupolaire électriques ; moment magnétique

Certains résultats sont beaucoup plus rapides à obtenir "microscopiquement". Par exemple pour un volume V contenant toutes les charges, la relation $\displaystyle\sum_i q_i \vec{v}_i = \frac{\mathrm{d}}{\mathrm{d}t} \sum_i q_i \vec{r}_i = \frac{\mathrm{d}}{\mathrm{d}t}\, \vec{p}$

($\vec{p}$ **moment dipolaire électrique**) est évidente, alors que la démonstration de son équivalent continu $\iiint_V \vec{j}\,(\vec{r},t)\,\mathrm{d}^3r = \dfrac{\mathrm{d}}{\mathrm{d}t}\iiint_V \rho\,(\vec{r},t)\,\vec{r}\,\mathrm{d}^3r$ fait appel à l'égalité

$$j_i = \vec{j}\cdot\overrightarrow{\mathrm{grad}}\,x_i = \mathrm{div}\,(x_i\vec{j}) - x_i\,\mathrm{div}\vec{j}$$

et à la L.C. $\partial_t\rho + \mathrm{div}\vec{j} = 0$. De même on établit facilement que la quantité $\iiint_V \vec{j}\,(\vec{r},t)\,(\vec{r}\cdot\hat{u})\,\mathrm{d}^3r$ s'écrit $\sum_i q_i\vec{v}_i\,(\vec{r}_i\cdot\hat{u}) = \dfrac{1}{2}\sum_i q_i(\vec{r}_i\wedge\vec{v}_i)\wedge\hat{u} + \dfrac{1}{2}\dfrac{\mathrm{d}}{\mathrm{d}t}\sum_i q_i\vec{r}_i\,(\vec{r}_i\cdot\hat{u})$, où $\vec{\mu} = \dfrac{1}{2}\sum_i q_i(\vec{r}_i\wedge\vec{v}_i)$ est un **moment magnétique**, et où $\sum_i q_i\vec{r}_i\,(\vec{r}_i\cdot\hat{u})$ est un terme (de type) **grandeur quadrupolaire**.

▷ Applications

1) Pour une distribution stationnaire (**magnétostatique**), on a

$$\overrightarrow{A}(\vec{r}_0) = \frac{\mu_0}{4\pi}\iiint_V \frac{\vec{j}(\vec{r})}{|\vec{r}-\vec{r}_0|}\,\mathrm{d}^3r = \frac{\mu_0}{4\pi r_0}\iiint_V \vec{j}(\vec{r})\left(1+\frac{\vec{r}\cdot\hat{u}_0}{r_0}+\cdots\right)\mathrm{d}^3r$$

à grande distance (avec $\hat{u}_0 = \dfrac{\vec{r}_0}{r_0}$) ; l'intégrale de $\vec{j}$, égale $\dot{\vec{p}}$, étant nulle ($\vec{p}$ indépendant de t) ainsi que la dérivée temporelle du terme quadrupolaire, la distribution de courants est équivalente à un moment magnétique :

$$\overrightarrow{A}(\vec{r}_0) \simeq \frac{\mu_0}{4\pi r_0^2}\,\vec{\mu}\wedge\hat{u}_0\ .$$

2) De même pour le **potentiel rayonné** (termes proportionnels à r_0^{-1}) par une distribution quelconque de charges et de courants, on a

$$\begin{aligned}\overrightarrow{A}(\vec{r}_0,t_0) &= \iiint\frac{\mu_0}{4\pi}\,\frac{\vec{j}\left(\vec{r},t-\frac{|\vec{r}-\vec{r}_0|}{c}\right)}{|\vec{r}-\vec{r}_0|}\,\mathrm{d}^3r\\ &\simeq \frac{\mu_0}{4\pi r_0}\left[\iiint_V\vec{j}\Big(\vec{r},t_0-\frac{r_0}{c}\Big)\,\mathrm{d}^3r + \frac{\mathrm{d}}{\mathrm{d}t}\iiint_V\vec{j}\Big(\vec{r},t_0-\frac{r_0}{c}\Big)\,\frac{\vec{r}\cdot\hat{u}_0}{c}\,\mathrm{d}^3r+\cdots\right]\end{aligned}$$

(développement au numérateur de $\dfrac{|\vec{r}-\vec{r}_0|}{c}$ pour r_0 grand). Avec les densités microscopiques on obtient immédiatement le début du développement multipolaire

$$\overrightarrow{A}(\vec{r}_0,t_0) \simeq \frac{\mu_0}{4\pi r_0}\left(\dot{\vec{p}} + \frac{1}{c}\,\dot{\vec{\mu}}\wedge\hat{u}_0 + \frac{1}{2c}\,\frac{\mathrm{d}^2}{\mathrm{d}t^2}\sum_i q_i\vec{r}_i\,(\vec{r}_i\cdot\hat{u}_0)+\cdots\right),$$

le second membre étant calculé à l'instant $t_0 - \dfrac{r_0}{c}$. Au premier ordre en r_0^{-1} l'action sur $\overrightarrow{A}$ de $\overrightarrow{\nabla}_{\vec{r}_0}$ équivaut à $-\dfrac{\hat{u}_0}{c}\,\partial_{t_0}$; on en déduit que $\overrightarrow{B} = \overrightarrow{\nabla}_{\vec{r}_0}\wedge\overrightarrow{A} = -\dfrac{\hat{u}_0}{c}\wedge\partial_{t_0}\overrightarrow{A}$ et, à partir de $\overrightarrow{\nabla}_{\vec{r}_0}\wedge\overrightarrow{B} = \epsilon_0\mu_0\,\partial_{t_0}\overrightarrow{E}$, que $\overrightarrow{E} = -c\,\hat{u}_0\wedge\overrightarrow{B}$ (comme pour une onde plane). Si $\partial_{t_0}\overrightarrow{A}$ est selon Oz, alors $\overrightarrow{E} = -\partial_{t_0}A_z\,\sin\theta\,\hat{\theta}$ (figure 50). Si $\vec{v}_i \ll c$, ou pour des courants de période T confinés dans une antenne de dimension $l \ll cT = \lambda$, le premier terme du développement est dominant : **rayonnement dipolaire électrique**.

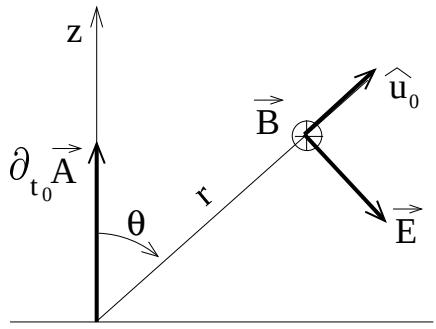


FIGURE 50

▷ Lissage des charges et des courants dans un milieu neutre

Regroupons les charges par paquets neutres α (molécules par exemple), et soient $\vec{r}_\alpha(t)$ le vecteur position de leur centre de masse et $\vec{r}_{i,\alpha}(t)$ le vecteur joignant le c.d.m. à une charge $q_{i,\alpha}$ d'un paquet ; pour chaque paquet posons $\vec{p}_\alpha(t) = \sum_{i\in\alpha} q_{i,\alpha}\vec{r}_{i,\alpha}(t)$ et $\vec{\mu}_\alpha(t) = \sum_{i\in\alpha} \frac{1}{2} q_{i,\alpha}\vec{r}_{i,\alpha}(t) \wedge \vec{v}_{i,\alpha}(t)$. 1) La forme "lissée" $\sum_\alpha \left(\sum_i q_{i,\alpha}\, f\left(\vec{r} - (\vec{r}_\alpha + \vec{r}_{i,\alpha})\right)\right)$ de ρ^{micro} (δ remplacé par f) s'écrit, à l'aide du développement au premier ordre $f\left(\vec{r} - (\vec{r}_\alpha + \vec{r}_{i,\alpha})\right) \simeq f(\vec{r} - \vec{r}_\alpha) - \vec{r}_{i,\alpha} \cdot \overrightarrow{\nabla} f(\vec{r} - \vec{r}_\alpha)$, et en tenant compte de la neutralité de chaque paquet : $-\sum_\alpha \vec{p}_\alpha(t) \cdot \overrightarrow{\nabla} f(\vec{r} - \vec{r}_\alpha(t)) = -\text{div}\left(\sum_\alpha \vec{p}_\alpha(t)\, f(\vec{r} - \vec{r}_\alpha(t))\right) = -\text{div}\, \vec{P}(\vec{r},t) = \rho_{\text{pol}}(\vec{r},t)$. $\overrightarrow{P}(\vec{r},t)$ représente une densité volumique de moments dipolaires. 2) De même un développement de la densité de courant microscopique "lissée" $\sum_\alpha \left(\sum_i q_{i,\alpha}\, (\vec{v}_\alpha + \vec{v}_{i,\alpha})\, f(\vec{r} - (\vec{r}_\alpha + \vec{r}_{i,\alpha}))\right) \simeq \sum_\alpha \left(-(\vec{p}_\alpha \cdot \overrightarrow{\nabla} f)\, \vec{v}_\alpha + \dot{\vec{p}}_\alpha\, f - \vec{\mu}_\alpha \wedge \overrightarrow{\nabla} f\right)$ (en négligeant le terme quadrupolaire) montre que la densité moyenne de courant s'écrit $\vec{j}(\vec{r},t) = \frac{\partial \overrightarrow{P}}{\partial t} + \overrightarrow{\text{rot}}\ \overrightarrow{M}$ avec $\overrightarrow{M}(\vec{r},t) = \sum_\alpha \vec{\mu}_\alpha(t)\, f(\vec{r} - \vec{r}_\alpha(t))$ si on peut négliger aussi $|\vec{p}_\alpha|\, |\vec{v}_\alpha|$ devant $|\vec{\mu}_\alpha|$.

REMARQUE. Dans les conducteurs (et pour des charges libres de même type), on écrit plus simplement

$$\sum_i q_i \vec{v}_i(t)\, f(\vec{r} - \vec{r}_i(t)) = \vec{v}(t) \sum_i q\, f(\vec{r} - \vec{r}_i(t)) = \rho(\vec{r},t)\, \vec{v}(\vec{r},t)\ ,$$

où ρ et $\vec{v}$ sont la densité de charge et la vitesse moyennes ; en régime variable $\rho\vec{v}$ correspond à $\partial_t \overrightarrow{P}$ dans la modélisation précédente.

Chapitre 8

Equations aux dérivées partielles ; propagation ; diffusion

En physique les équations aux dérivées partielles (**E.D.P.**) relient des champs, dépendant en général de l'espace et du temps $f(\vec{r}, t)$, à leurs dérivées partielles. Lorsque la variable t est présente il est naturel, comme le montre l'analogie entre corde vibrante et chaîne d'oscillateurs couplés, de considérer les E.D.P. comme des systèmes dynamiques à un nombre infini (continu) de degrés de liberté indicés par $\vec{r}$. Cependant un aspect nouveau très important est le rôle joué par les **conditions aux limites (C.L.)**, par exemple la donnée de f à tout instant sur des surfaces, à la place des (ou avec les) **conditions initiales (C.I.)**. On étudiera essentiellement le cas particulier, mais non trivial, des E.D.P. linéaires à coefficients constants (sans ou avec second membre), notamment celles introduites à propos des cordes vibrantes (d'Alembert 1747), de la conduction thermique (Fourier 1822), de la gravitation et de l'électromagnétisme (Laplace $\simeq$1780, Poisson 1813, Maxwell 1864, *etc.*) et de la mécanique quantique (Schrödinger 1925). Quelques exemples autres (non linéaires en particulier) montreront au lecteur que la richesse du domaine des E.D.P. ne se résume pas à l'étude de ces exemples historiques.

8.1 CHAINES DE SYSTÈMES DYNAMIQUES COUPLÉS ; LIMITE CONTINUE

On étudie, à partir d'exemples, quelques systèmes infinis d'E.D. linéaires et stationnaires (pour des grandeurs notées $\Psi_n(t) \equiv \Psi(n,t)$; $n = -\infty \cdots -1, 0, 1 \cdots \infty$), présentant une invariance de translation discrète ($\Psi_n \to \Psi_{n+1}$). Dans la limite continue les $\Psi(n,t)$ deviennent une onde $\Psi(x,t)$ et le système infini d'E.D. une E.D.P. ; de nombreuses grandeurs passent ainsi du "discret" au "continu".

8.1.1 Chaines d'oscillateurs ; rôle des conditions aux limites ; phonons

▷ Petites oscillations longitudinales de pendules simples couplés (figure 1)

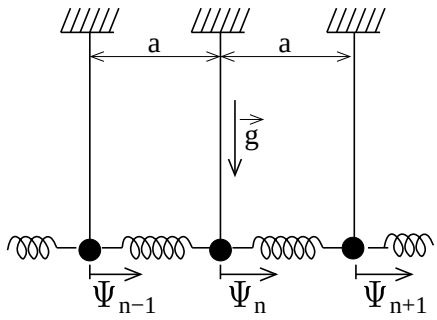


FIGURE 1

Les écarts aux positions d'équilibre $\Psi_n(t)$ satisfont les équations :

$m\ddot{\Psi}_n = -m\omega_0^2\Psi_n - K(\Psi_n - \Psi_{n-1}) - K(\Psi_n - \Psi_{n+1})$ (ω_0 pulsation d'un pendule isolé) .

Comme pour un nombre fini d'oscillateurs, il est naturel de chercher les solutions sous forme de modes (*cf.* section 6.2.2). En posant $\Psi_n(t) = A_n\, e^{-i\omega t}$, on obtient la relation de récurrence d'ordre 2 :

$$\gamma A_{n-1} + (\omega^2 - \omega_0^2 - 2\gamma)\, A_n + \gamma\, A_{n+1} = 0 \qquad \left(\gamma = \frac{K}{m}\right) .$$

(Si on pose $\Psi_n(t) = a_n\, \cos(\omega t + \Phi)$ les a_n satisfont la même relation.) Sa solution générale dépend à priori de deux constantes ; on vérifie facilement, en essayant des solutions $A_n = \alpha r^n$, qu'elle s'écrit $\alpha_1 r_1^n + \alpha_2 r_2^n$ où r_1 et r_2 sont les racines de l'équation du second degré $\gamma r^2 + (\omega^2 - \omega_0^2 - 2\gamma)\, r + \gamma = 0$. r_1 et r_2 sont de la forme $r_{1,2} = e^{\pm\varphi}$ si $\omega^2 \in [0, \omega_0^2]$ (solutions qui disparaissent si $\omega_0 = 0$), $r_{1,2} = e^{\pm i\varphi}$ si $\omega^2 \in [\omega_0^2, \omega_0^2 + 4\gamma]$ et $r_{1,2} = -e^{\pm\varphi}$ si $\omega^2 > \omega_0^2 + 4\gamma$. Ce passage des E.D. à une relation de récurrence puis aux **solutions** $\underline{\Psi}_n(t) = \underline{\Psi}\, r^n\, e^{-i\omega t}$, où r satisfait une équation algébrique, s'applique à de nombreux autres exemples.

▷ Rôle des conditions aux limites (C.L.)

Pour un **système infini** les solutions acceptables sont celles avec $|r| = 1$ car elles doivent rester bornées quand $n \to \pm\infty$. Bien que la périodicité spatiale "a" du sytème (*cf.* figure 1) n'apparaisse pas dans les équations, on pose $r = e^{\pm ika}$; k est défini modulo $\dfrac{2\pi}{a}$ et en pratique on choisit $k \in \left]-\dfrac{\pi}{a}, \dfrac{\pi}{a}\right]$. La solution générale (complexe) $\Psi_n(t) = \alpha_1\, e^{-i(\omega t - nka)} + \alpha_2\, e^{-i(\omega t + nka)}$ s'interprète alors comme une somme de deux ondes propagatives se propageant en sens inverses le long de la chaîne de pendules. De $\gamma r + (\omega^2 - \omega_0^2 - 2\gamma) + \gamma r^{-1} = 0$ on déduit la relation $\omega^2 = \omega_0^2 + 2\gamma\,(1 - \cos ka)$ entre ω et k appelée **relation de dispersion des ondes** (figure 2). Cette relation est également valable pour les solutions (réelles) $\Psi_n(t) = a\, \cos(\omega t + \Phi)\, \cos(nka + \varphi)$. Ces ondes stationnaires sont l'analogue des modes propres étudiés à la section 6.2.2, les fréquences propres (discrètes) étant ici remplacées par une bande (continue) de fréquences.

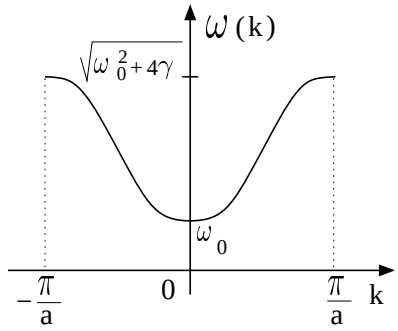


FIGURE 2

Pour un **système semi-infini** ($n = 0, 1 \cdots \infty$) la racine r telle que $|r| < 1$ est aussi acceptable. En posant $r = e^{-|k|a}$ (si $\omega^2 < \omega_0^2$) ou $r = -e^{-|k|a}$ (si $\omega^2 > \omega_0^2 + 4\gamma$), on obtient les solutions de la forme $e^{-i\omega t}\, e^{-n|k|a}$ ou $e^{-i\omega t}\, (-1)^n\, e^{-n|k|a}$, et les relations $\omega^2 = \omega_0^2 + 2\gamma\,(1 \mp \cosh|k|a)$. De telles solutions non propagatives sont observées si on impose au pendule "0" une pulsation ω satisfaisant les conditions ci-dessus.

Le cas d'un **système fini de N pendules identiques** peut être déduit de celui du système infini en imposant à ce dernier des C.L. particulières. Par exemple toute solution du système infini de la figure 1 qui vérifie $\Psi_0 = \Psi_{N+1} = 0$ est acceptable pour le système des N pendules de la figure 3a. Appliquées à $A_n \propto \cos(nka + \varphi)$, on trouve facilement que ces C.L. conduisent à N modes propres $\Psi_n^{(p)}(t) \propto \cos(\omega_p t + \Phi) \sin nk_p a$ avec $(N+1)\, k_p a = p\pi$ ($p = 1, 2 \cdots N$) et $\omega_p = \omega(k_p)$. Si le pendule N est libre (figure 4a), les C.L. à imposer au système infini sont $\Psi_0 = 0$ et $\Psi_N = \Psi_{N+1}$ (ainsi le pendule N n'est soumis à aucune force de la part du ressort "de droite") ; on a alors $\left(N + \dfrac{1}{2}\right) k_p a = \left(p - \dfrac{1}{2}\right) \pi$ ($p = 1, 2 \cdots N$). Les figures 3b et 4b montrent les deux premiers modes propres de ces systèmes (pour $N = 5$).

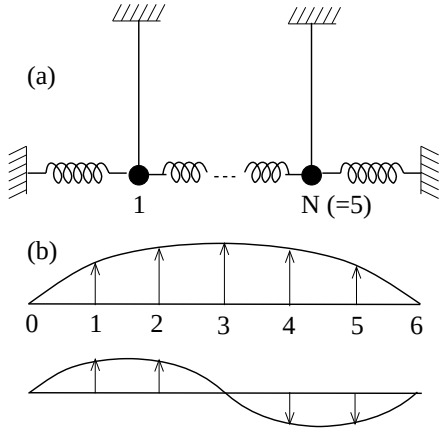


FIGURE 3

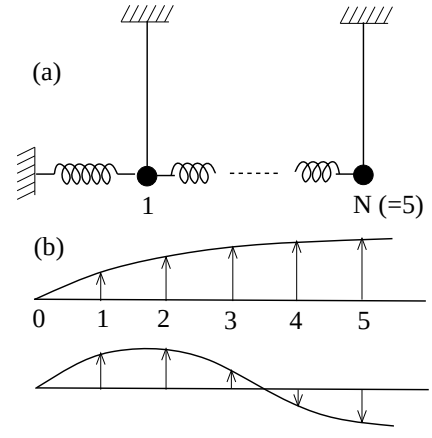


FIGURE 4

AUTRES EXEMPLES. **Petites oscillations longitudinales d'une chaîne d'atomes** (cristal à une dimension)

1) Dans le cas où les atomes sont identiques et les interactions ne se limitent pas aux plus proches voisins (figure 5a) on a :

$$m\,\ddot{\Psi}_n = -\sum_{n'=-\infty}^{\infty} K_{n-n'}\Psi_{n'} \quad \text{avec } K_q = K_{-q} \text{ et } \sum_{q=-\infty}^{\infty} K_q = 0\ .$$

Ces équations se déduisent de l'existence d'une énergie potentielle d'interaction quadratique invariante par translation $E_p = \frac{1}{2}\sum_{n,n'} K_{n-n'}\Psi_n\Psi_{n'}$. La relation de dispersion pour des ondes (**phonons**) obtenue en posant $\Psi_n = \Psi e^{-i(\omega t - nka)}$ est $\omega^2 = m^{-1}\sum_{q=-\infty}^{\infty} K_q e^{iqka}$.

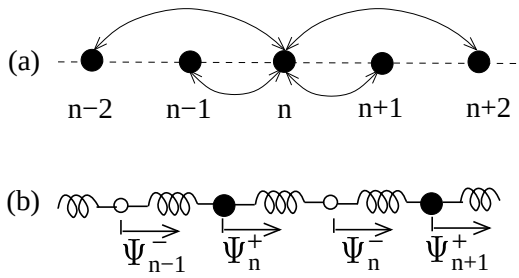


FIGURE 5

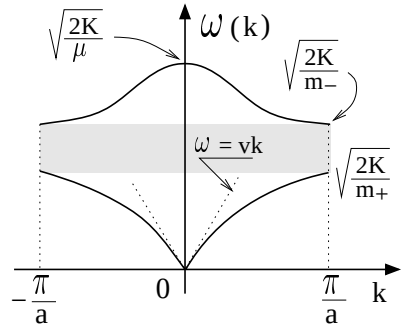


FIGURE 6

2) Si la chaîne est la répétition d'un motif de deux atomes (figure 5b) et si (pour simplifier) les interactions se font entre plus proches voisins, on a :

$$m_+\ddot{\Psi}^+_n = -K(\Psi^+_n - \Psi^-_n) - K(\Psi^+_n - \Psi^-_{n-1}) \quad \text{et} \quad m_-\ddot{\Psi}^-_n = -K(\Psi^-_n - \Psi^+_n) - K(\Psi^-_n - \Psi^+_{n+1})\ .$$

En posant $\Psi^\pm_n = \Psi^\pm e^{-i(\omega t - nka)}$, il vient :

$$(m_+\omega^2 - 2K)\,\Psi^+ + K\,(1 + e^{-ika})\,\Psi^- = 0 \ \text{ et } \ (m_-\omega^2 - 2K)\,\Psi^- + K\,(1 + e^{ika})\,\Psi^+ = 0\ .$$

Ce système de deux équations à deux inconnues n'a de solutions $\Psi^\pm \neq 0$ que si (petit calcul) :

$$\omega^2 = K\,(m_+^{-1} + m_-^{-1}) \pm K\,\big((m_+^{-1} + m_-^{-1})^2 - 4m_+^{-1}m_-^{-1}\,\sin^2\tfrac{ka}{2}\big)^{\frac{1}{2}}\ .$$

Dans ce cas le graphe de la relation de dispersion des phonons (figure 6) contient deux branches : l'une dite "sonore" pour laquelle $\omega^2 \simeq v^2k^2$ et $\Psi^+ \simeq \Psi^-$ pour $k \simeq 0$ (avec $2(m_+ + m_-)\,v^2 = Ka^2$) ; l'autre dite "optique" pour laquelle $\omega^2 \simeq 2K\,\mu^{-1}$ et $m_+\Psi^+ + m_-\Psi^- \simeq 0$ pour $k \simeq 0$ (avec μ masse réduite).

8.1.2 Limite continue ; cordes vibrantes ; lignes électriques ; hydrodynamique ; impédances

Dans cette section on considère les équations (analogues à celles des pendules avec $\omega_0 = 0$)

$$m\,\ddot{\Psi}_n = -\frac{T}{a}(\Psi_n - \Psi_{n-1}) - \frac{T}{a}(\Psi_n - \Psi_{n+1}) = F_n - F_{n+1}\ .$$

La relation de dispersion associée est $\omega^2 = \dfrac{2T}{ma}\,(1 - \cos ka)$. Elles décrivent les **oscilla-**

tions transversales d'une chaîne de masses couplées par des ressorts de tension T (figure 7a) dans l'approximation où les lignes des ressorts font des petits angles avec l'axe "x" de la chaîne. (F_n est la composante transverse de la force qu'exerce le ressort "de gauche" sur la masse n.) Des équations identiques décrivent la **chaîne de quadripoles électriques** de la figure 7b, moyennant les correspondances $m \leftrightarrow L$, $\dfrac{T}{a} \leftrightarrow C^{-1}$, $\dot{\Psi}_n \leftrightarrow I_n$ et $F_n \leftrightarrow V_n$.

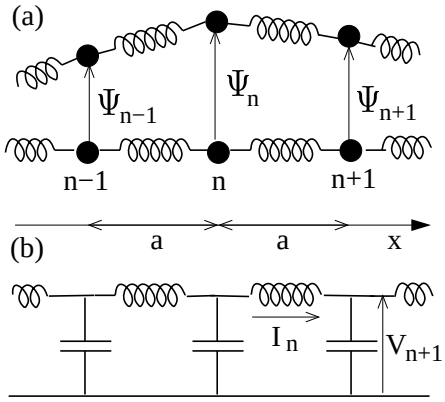


FIGURE 7

▷ Limite continue

L'approximation continue (ou "grande échelle") correspond au cas où, pour tout n, $\Psi_n(t)$ et $\Psi_{n+1}(t)$ sont proches. Les $\Psi_n(t)$ sont alors considérés comme les valeurs en $x = na$ d'une fonction régulière $\Psi(x,t)$ à laquelle on peut appliquer le développement :

$$\Psi_{n\pm1}(t) = \Psi(x \pm a, t) = \Psi(x,t) \pm a\,\partial_x\Psi(x,t) + \frac{a^2}{2}\,\partial_x^2\,\Psi(x,t)\cdots$$

Dans l'exemple de la figure 7a le système d'équations pour les Ψ_n devient (après division par a) :

$$\rho\,\frac{\partial^2\Psi}{\partial t^2} = T\,\frac{\partial^2\Psi}{\partial x^2} + \cdots = -\frac{\partial F}{\partial x} + \cdots \qquad \left(\rho = \frac{m}{a}\ ,\ F(x) = -T\frac{\partial\Psi}{\partial x}\right)\ .$$

Les termes oubliés (dont le premier est $\dfrac{a^2}{3}\,T\,\partial_x^4\Psi$) sont négligeables (par rapport à $T\,\partial_x^2\Psi$) si l'échelle de longueur typique L des variations spatiales de $\Psi(x,t)$ vérifie $L \gg a$, et ils deviennent nuls dans la limite continue $a \to 0$, $m \to 0$ avec $\dfrac{m}{a} = \rho$ masse linéique finie. L'équation obtenue est celle d'une **corde vibrante** qu'on peut aussi obtenir directement

en écrivant (figure 8) :

$$(\rho\,\mathrm{d}x)\,\partial_t^2\Psi(x,t) = F(x,t) - F(x+\mathrm{d}x,t) \quad \text{avec} \quad F(x,t) = -T\,\partial_x\Psi(x,t)\ .$$

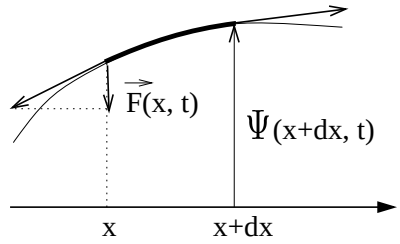


FIGURE 8

La vitesse de propagation des ondes est $v = \sqrt{\dfrac{T}{\rho}}$. (Elle est obtenue aussi à partir de la limite continue $\omega^2 = \dfrac{2T}{ma}\,(1-\cos ka) \to \dfrac{T}{\rho}\,k^2$.)

REMARQUES. 1) Dans l'exemple de la figure 7b les équations $V_n = L\,\dfrac{\mathrm{d}I_n}{\mathrm{d}t} + V_{n+1}$ et $C\,\dfrac{\mathrm{d}V_n}{\mathrm{d}t} = I_{n-1} - I_n$ deviennent, dans la limite continue, $l\,\partial_t I = -\partial_x V$ et $\gamma\,\partial_t V = -\partial_x I$, qui sont celles d'une **ligne électrique** d'inductance et de capacité linéiques $l = \dfrac{L}{a}$ et $\gamma = \dfrac{C}{a}$. Elles conduisent pour I à l'E.D.P. $\gamma l\,\partial_t^2 I = \partial_x^2 I$ qui est aussi celle vérifiée par V ou $Q = \int^t I(x,t')\,\mathrm{d}t'$ (analogue électrique de Ψ) ; ici $v = (\gamma l)^{-\frac{1}{2}}$.

2) Dans l'exemple de la figure 1, la limite continue $a \to 0$, $\frac{m}{a} \to \rho$ et $Ka \to E$ conduit à l'E.D.P.

$$\partial_t^2\psi = -\omega_0^2\psi + v^2\,\partial_x^2\psi \quad \text{avec} \quad v = \sqrt{\tfrac{E}{\rho}}\ ;$$

la relation de dispersion est $\omega^2 = \omega_0^2 + v^2k^2$. Pour $\omega_0 = 0$ cette E.D.P. décrit les ondes longitudinales dans une barre dont E est le **module d'Young**.

▷ Bilans énergétiques

En multipliant l'équation pour Ψ_n par $\dot\Psi_n$, on obtient le bilan énergétique pour le système ouvert constitué par la $n^{\text{ième}}$ masse et le ressort "de droite" :

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\,m\,\dot\Psi_n^2 + \frac{1}{2}\,\frac{T}{a}\,(\Psi_n - \Psi_{n+1})^2\right) = F_n\,\dot\Psi_n - F_{n+1}\,\dot\Psi_{n+1}\ .$$

($\dfrac{1}{2}\,\dfrac{T}{a}\,(\Psi_n - \Psi_{n+1})^2 = T\,\delta l_n$ est le surcroît d'énergie du ressort lié à son allongement.) Ce bilan, divisé par a, conduit, dans la limite continue à la L.C. de l'énergie (qu'on peut aussi obtenir en multipliant l'E.D.P. par $\partial_t\Psi$) :

$$\partial_t e + \partial_x j = 0 \quad \text{avec } e = \frac{1}{2}\rho\,(\partial_t\Psi)^2 + \frac{1}{2}T\,(\partial_x\Psi)^2 \ \text{ et } \ j = -T\,(\partial_x\Psi)(\partial_t\Psi)\ .$$

e est une **densité linéique d'énergie** (cinétique + potentielle liée à l'allongement de la corde) et $j = F\,\partial_t\Psi$ un **flux d'énergie** (ou courant). Le lecteur établira (et interprètera) un bilan analogue pour la ligne électrique ; on a alors $j = VI = -\gamma^{-1}\,(\partial_x Q)(\partial_t Q)$.

Comme pour les oscillateurs, l'E.D.P. $\rho\,\partial_t^2\Psi - T\,\partial_x^2\psi = 0$ peut se déduire, inversement, de la L.C. de l'énergie. Par exemple, pour une corde fixée à ses deux extrémités ($\Psi(0,t) = \Psi(L,t) = 0$), on écrit

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\int_0^L \left[\frac{1}{2}\rho\,(\partial_t\Psi)^2 + \frac{1}{2}T\,(\partial_x\Psi)^2\right]\mathrm{d}x = \int_0^L \partial_t\Psi\left[\rho\,\partial_t^2\Psi - T\,\partial_x^2\Psi\right]\mathrm{d}x$$

(après intégration par parties de $T\,(\partial_x\Psi)(\partial_x\partial_t\Psi)$). L'E.D.P. résulte alors de ce que $\partial_t\Psi$ peut être, à un instant donné, une C.I. arbitraire.

REMARQUE : cette méthode permet de modéliser simplement une tige "rigide" (**corde de piano** ; figure 9) en ajoutant à la densité linéique d'énergie un **terme de courbure** $\frac{1}{2}D\left(\frac{\partial^2\Psi}{\partial x^2}\right)^2$.

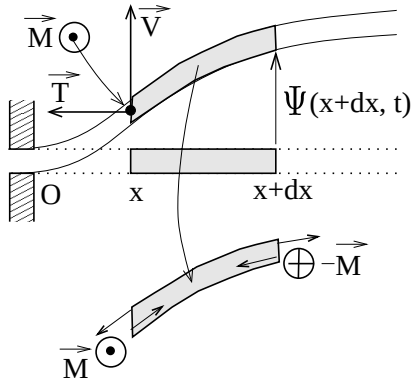


FIGURE 9

Du terme supplémentaire dans la dérivée temporelle de l'énergie entre 0 et L

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^L \frac{D}{2}\left(\frac{\partial^2\Psi}{\partial x^2}\right)^2\mathrm{d}x = \int_0^L D\,(\partial_x^2\Psi)(\partial_x^2\partial_t\Psi)\,\mathrm{d}x = \int_0^L D\,(\partial_t\Psi)(\partial_x^4\Psi)\,\mathrm{d}x$$

(en tenant compte des C.L. $\partial_x\Psi(0,t) = \partial_x\Psi(L,t) = 0$), on déduit l'E.D.P. (valable aussi pour des oscillations de plaques minces) :

$$\rho\frac{\partial^2\Psi}{\partial t^2} = T\,\frac{\partial^2\Psi}{\partial x^2} - D\,\frac{\partial^4\Psi}{\partial x^4}\ .$$

(Une description en termes de forces et de couples est plus subtile ; il faut écrire les équations $T(x,t) = T(x+\mathrm{d}x,t)$ et $\rho\,\partial_t^2\Psi(x,t) = V(x,t) - V(x+\mathrm{d}x,t)$ pour les forces, $M(x,t) - M(x+\mathrm{d}x,t) - T\,\partial_x\Psi(x,t)\,\mathrm{d}x - V(x,t)\,\mathrm{d}x = 0$ pour les couples, et faire appel à la relation $M = -D\,\partial_x^2\Psi$ entre couple et courbure.)

▷ **Impédance** (*cf.* aussi section 6.4.3)

La notion d'impédance pour une onde est importante en physique lorsqu'on s'intéresse à des transferts d'énergie de l'onde vers d'autres systèmes (en régime sinusoïdal).
Pour une ligne électrique, si

$$I(x,t) = a\ (\exp i(\omega t - kx) + r_I\ \exp i(\omega t + kx))$$

(solution générale de l'E.D.P. pour ω fixé), on déduit de $\gamma\,\partial_t V = -\partial_x I$ et $v = \frac{\omega}{k} = (\gamma l)^{-\frac{1}{2}}$ que

$$V(x,t) = Z_0 a\ (\exp i(\omega t - kx) + r_V\ \exp i(\omega t + kx))$$

avec $Z_0 = \sqrt{\frac{l}{\gamma}}$ et $r_V = -r_I$ (**coefficients de réflexion**). Z_0 qui représente le rapport $\frac{V}{I}$ lorsqu'il n'y a pas d'onde réfléchie ($r_V = r_I = 0$ "adaptation d'impédance") est l'**impédance caractéristique** de la ligne (limite continue de l'impédance itérative des quadripoles). Le rapport

$$Z(x) = \tfrac{V(x,t)}{I(x,t)}$$

est l'impédance en x. On déduit de ces définitions que si une impédance Z est branchée au bout de la ligne (pris comme origine $x = 0$), on a $Z = Z_0\,\frac{1-r_I}{1+r_I}$ d'où $r_I = \frac{Z_0 - Z}{Z_0 + Z} = -r_V$. Un calcul élémentaire donne alors le flux moyen d'énergie sur la ligne : $\frac{1}{2}\,\Re\mathrm{e}(V\overline{I}) = \frac{1}{2}\,Z_0\,|a|^2(1 - |r_I|^2)$. Ce flux est aussi égal, pour $z = 0$, à : $\frac{1}{2}\,\Re\mathrm{e}Z\,|I(0,t)|^2$ (puissance dissipée dans Z). Un calcul simple montre que l'impédance vue sur la ligne en $z = -L$ est :

$$Z(-L) = Z_0\,\tfrac{Z\,\cos kL + iZ_0\,\sin kL}{Z_0\,\cos kL + iZ\,\sin kL}\ .$$

Si pour $z > 0$ on a une ligne d'impédance Z_2, et pour $z < -L$ une ligne d'impédance $Z_1 \neq Z_2$, l'**adaptation** des deux lignes (passage intégral de l'énergie de 1 à 2) s'obtient lorsque $Z(-L) = Z_1$, c'est à dire $\cos kL = 0$ (ligne intermédiaire quart d'onde) et $Z_0^2 = Z_1 Z_2$.
Cet exemple classique a son analogue pour la corde vibrante qu'on déduit des correspondances $V \leftrightarrow F = -T\,\partial_x\Psi$, $I \leftrightarrow \partial_t\Psi$, $l \leftrightarrow \rho$ et $\gamma^{-1} \leftrightarrow T$; on a alors $Z_0 = \sqrt{\rho T}$. Si on place au bout de la corde ($x = 0$) le système mécanique de la figure 10 (masse m attachée à un ressort de raideur K et coulissant avec frottement), l'équation $m\,\partial_t^2\Psi = -K\Psi - f\,\partial_t\Psi - T\,\partial_x\Psi$ montre que l'**impédance mécanique** $\frac{-T\,\partial_x\Psi}{\partial_t\Psi}$ à ce bout n'est autre que $Z = f + i\big(m\omega - \frac{K}{\omega}\big)$, analogue mécanique de $Z = R + i\big(L\omega - \frac{1}{C\omega}\big)$.

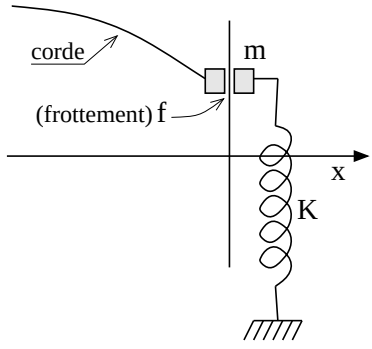


FIGURE 10

▷ Exemples en hydrodynamique

1) Les **ondes sonores dans un tuyau à paroi extensible** (figure 11) sont régies par les équations :

$$\rho\,\partial_t^2\Psi = -\partial_x P' \quad \text{et} \quad -\chi\,P' = \tfrac{\delta V}{V} = \partial_x\Psi + DP' \quad (\text{donc} \quad P' = -(\chi+D)^{-1}\,\partial_x\Psi)\ .$$

P' est la surpression, D le coefficient d'extensibilité défini par $\frac{\mathrm{d}S}{S} = DP'$ (S section du tuyau) et $\frac{\delta V}{V}$ la variation relative de volume d'une tranche de fluide de masse volumique ρ et de compressibilité χ ; on vérifie la L.C.

$$\partial_t e + \partial_x j = 0 \quad \text{avec} \quad e = \tfrac{1}{2}\rho S(\partial_t\Psi)^2 + \tfrac{1}{2}(\chi+D)SP'^2 \quad \text{et} \quad j = P'S(\partial_t\Psi)\ .$$

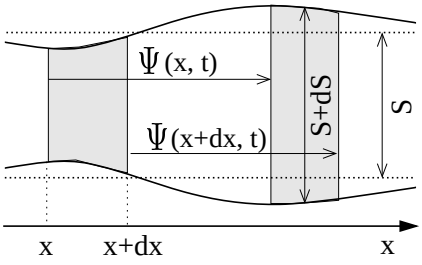


FIGURE 11

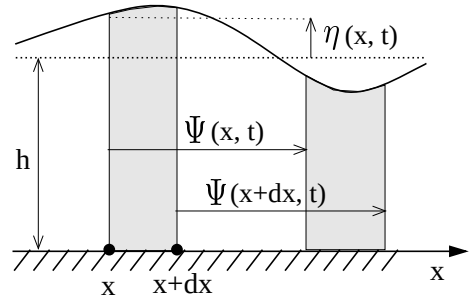


FIGURE 12

2) Les **ondes de surface dans un canal** peu profond de hauteur h et section S (figure 12) sont régies par les équations analogues (avec $\chi = 0$) :

$$\rho\,\partial_t^2\Psi = -\partial_x P' \ \text{ avec } \ P' = \rho g\eta \ \text{ et } \ \tfrac{\delta V}{V} = \partial_x\Psi + \tfrac{\eta}{h} = 0 \ \text{(conservation du volume)}\ .$$

Dans ces exemples le rôle de F est tenu par P' et Ψ est remplacé par le volume déplacé $S\Psi$; on a alors $Z_0 = \frac{\rho v}{S}$ avec $v = (\rho(\chi+D))^{-\frac{1}{2}}$ ou $v = (gh)^{\frac{1}{2}}$, et $Z = \frac{P'}{S(\partial_t\Psi)}$ est l'**impédance acoustique**.

8.2 SOLUTIONS DE QUELQUES E.D.P. DYNAMIQUES

8.2.1 E.D.P. linéaires à coefficients constants et solutions ondes planes

Ces E.D.P. (sans second membre, c.a.d. en l'absence de sources) concernent des champs dans des milieux dont les propriétés sont les mêmes en tout point (homogénéité) et à tout instant (stationnarité), et tels que ces champs peuvent se superposer (linéarité). Elles possèdent des solutions onde plane en fonction desquelles toute solution peut s'exprimer.

▷ Onde plane $\underline{A}\,e^{-i(\omega t - \vec{k}\cdot\vec{r})}$ (notation complexe)

Comme on le verra ω et $\vec{k}$ peuvent être complexes : $\omega = \omega_1 + i\omega_2$ et $\vec{k} = \vec{k}_1 + i\vec{k}_2$. Pour l'onde physique (réelle) ω_1 et $\vec{k}_1$ sont liés à la **propagation** tandis que ω_2 et $\vec{k}_2$ décrivent la **(dé)croissance de l'amplitude**. Par exemple pour une grandeur scalaire $\Re\mathrm{e}\big[A\,e^{-i(\omega t-\vec{k}\cdot\vec{r})}\big] = a\,e^{\omega_2 t}e^{-\vec{k}_2\cdot\vec{r}}\,\cos(\omega_1 t - \vec{k}_1\cdot\vec{r}_1 + \varphi)$ (avec $A = a\,e^{-i\varphi}$). Pour une grandeur vectorielle (*cf.* exemple à la section 2.5.3), $\overrightarrow{A}$ décrit les états de polarisation (rectiligne si $\overrightarrow{A}$ est réel). Sur une onde plane les dérivées partielles agissent comme des multiplications :

$\partial_t \equiv -i\omega$, $\partial_x \equiv ik_x \cdots$. Par exemple pour $\overrightarrow{E}(\vec{r},t) = \overrightarrow{E}_0 \exp -i(\omega t - \vec{k}\cdot\vec{r})$ on a $\partial_t^2 \overrightarrow{E} = -\omega^2 \overrightarrow{E}$, $\operatorname{div}\overrightarrow{E} = i\vec{k}\cdot\overrightarrow{E}$, $\overrightarrow{\operatorname{rot}}\,\overrightarrow{E} = i\vec{k}\wedge\overrightarrow{E}$, $\Delta\overrightarrow{E} = -k^2\overrightarrow{E}$, *etc.*

▷ Relation de dispersion

Lorsqu'on cherche des solutions sous la forme d'ondes planes, le système d'E.D.P. devient un système d'équations algébriques linéaires sans second membre pour les amplitudes des ondes. Comme un tel système n'admet de solutions non nulles que si il existe des relations entre les coefficients, lesquels dépendent ici de ω et $\vec{k}$, on en déduit que les ondes planes sont solutions si il existe une (au moins) relation entre ω et $\vec{k}$. Donnons des exemples. Pour l' **E.D.P. de diffusion** et l'E.D.P. de propagation on obtient

$$\partial_t f = D\,\Delta f \Longrightarrow \omega = -ik^2 D \quad , \quad \partial_t^2 f - v^2\,\Delta f = 0 \Longrightarrow \omega^2 = v^2 k^2 \ ,$$

et pour un **milieu diélectrique conducteur** de conductivité γ (*cf.* cours d'électricité) :

$$\Delta\overrightarrow{E} = \mu\left(\epsilon\,\partial_t^2\overrightarrow{E} + \gamma\,\partial_t\overrightarrow{E}\right) \Longrightarrow k^2 = \epsilon\mu\omega^2 + i\mu\gamma\omega \ .$$

Dans ce dernier exemple, dont les deux premiers sont les cas limites $\omega \ll \dfrac{\gamma}{\epsilon}$ et $\omega \gg \dfrac{\gamma}{\epsilon}$, l'"imaginaire pur" i apparaît dans la relation ; ceci est lié à la présence de processus irréversibles (ici la conduction ohmique $\vec{j} = \gamma\overrightarrow{E}$) ; comme $2\vec{k}_1\cdot\vec{k}_2 = \Im m\vec{k}^2 = \mu\gamma\omega > 0$, l'onde s'atténue au cours de sa propagation (on suppose ω réel positif).

REMARQUE : l'**E.D.P. de Schrödinger** $i\hbar\partial_t\psi = -\dfrac{\hbar^2}{2m}\,\Delta\psi$, identique formellement à l'équation de diffusion avec $D = \dfrac{i\hbar}{2m}$, mais réversible (laissée invariante par les changements $t \to -t$ et $\psi \to \overline{\psi}$), conduit à la relation réelle : $\omega = \dfrac{\hbar k^2}{2m}$. (*cf.* aussi les E.D.P. relativistes section 4.4.5.)

Pour l'**E.D.P. des ondes élastiques** en milieu isotrope

$$\rho\,\partial_t^2\overrightarrow{\Psi} = \mu\,\Delta\overrightarrow{\Psi} + \left(\chi^{-1} + \frac{\mu}{3}\right)\overrightarrow{\operatorname{grad}}(\operatorname{div}\overrightarrow{\Psi}) \ ,$$

on obtient $(\rho\omega^2 - \mu k^2)\,\overrightarrow{\Psi} = \left(\chi^{-1} + \dfrac{\mu}{3}\right)(\vec{k}\cdot\overrightarrow{\Psi})\,\vec{k}$ ou, en séparant les composantes parallèle et perpendiculaire à $\vec{k}$: $\left(\rho\omega^2 - \left(\chi^{-1} + \dfrac{4}{3}\mu\right)k^2\right)\,\Psi_{\parallel} = 0$ et $(\rho\omega^2 - \mu k^2)\,\overrightarrow{\Psi}_{\perp} = 0$. Ce système a deux solutions ; soit $\rho\omega^2 = \left(\chi^{-1} + \dfrac{4}{3}\mu\right)k^2$ et $\overrightarrow{\Psi}_{\perp} = 0$ (onde sismique longitudinale), soit $\rho\omega^2 = \mu k^2$ et $\Psi_{\parallel} = 0$ (onde sismique transversale) ; à chacune est associée une polarisation.

Milieux diélectriques anisotropes parfaits. Des équations de Maxwell (*cf.* section 7.5.1) $\operatorname{div}\overrightarrow{D} = 0$, $\operatorname{div}\overrightarrow{B} = 0$, $\overrightarrow{\operatorname{rot}}\,\overrightarrow{E} = -\partial_t\overrightarrow{B}$ et $\overrightarrow{\operatorname{rot}}\,\overrightarrow{H} = \partial_t\overrightarrow{D}$ avec $\overrightarrow{B} = \mu_0\overrightarrow{H}$ et $\overrightarrow{D} = [\boldsymbol{\epsilon}]\,\overrightarrow{E}$ ($[\boldsymbol{\epsilon}]$ matrice symétrique réelle), on déduit les relations :

$$\vec{k}\cdot\overrightarrow{D} = 0 \ , \ \vec{k}\cdot\overrightarrow{B} = 0 \ , \ \vec{k}\wedge\overrightarrow{E} = \omega\overrightarrow{B} \ \text{ et } \ \vec{k}\wedge\overrightarrow{H} = -\omega\overrightarrow{D} \ .$$

Elles permettent immédiatement de déterminer les directions relatives des vecteurs $\vec{k}$, $\overrightarrow{D}$, $\overrightarrow{H}$, $\overrightarrow{E}$ et $\overrightarrow{\mathcal{P}} = \overrightarrow{E} \wedge \overrightarrow{H}$ pour une onde plane (figure 13) et, en écrivant $\vec{k} \wedge (\vec{k} \wedge \overrightarrow{E})$, d'obtenir la relation :

$$k^2 \overrightarrow{E}_{\perp} = \mu_0 \omega^2 [\boldsymbol{\epsilon}] \overrightarrow{E} \ .$$

Celle-ci n'est aisément exploitable que dans une base où $[\boldsymbol{\epsilon}]$ est diagonale ($D_x = \epsilon_0 n_x^2 E_x \cdots$), et si on choisit $\vec{k}$ dans un plan de symétrie, par exemple $\vec{k} = (k_x, 0, k_z)$. Alors les équations pour E_y et pour le couple (E_x, E_z) se séparent. De celle pour E_y on déduit $k^2 = k_x^2 + k_z^2 = \epsilon_0 \mu_0 \omega^2 n_y^2$ et $\overrightarrow{D}$ parallèle à $\hat{y}$. De celles pour (E_x, E_z) il résulte (petit calcul) que $\dfrac{k_x^2}{n_z^2} + \dfrac{k_z^2}{n_x^2} = \epsilon_0 \mu_0 \omega^2$ et que $\overrightarrow{D}$ est perpendiculaire à $\vec{k}$ et $\hat{y}$. Dans le cas particulier d'un **cristal uniaxe** où $n_x = n_y = n_o \neq n_z = n_e$, la symétrie de révolution permet toujours de choisir l'axe "x" de sorte que $k_y = 0$. La **surface des vecteurs d'onde** (lieu de l'extrémité de $\vec{k}$ à ω fixé) est alors constituée, d'une part par la sphère $k^2 = \epsilon_0 \mu_0 \omega^2 n_o^2$ pour les **ondes "ordinaires"** (car analogues à celles d'un milieu isotrope avec en particulier $\overrightarrow{D} \parallel \overrightarrow{E}$ et $\overrightarrow{\mathcal{P}} \parallel \vec{k}$), et d'autre part par l'ellipsoïde $\dfrac{k_z^2}{n_o^2} + \dfrac{k_x^2 + k_y^2}{n_e^2} = \epsilon_0 \mu_0 \omega^2$ pour les **ondes "extraordinaires"**. La figure 14 représente ces surfaces en coupe et, pour une direction donnée de $\vec{k}$, les caractéristiques des deux ondes. Pour l'onde extraordinaire la relation $k_x D_x + k_z D_z = 0$ équivalente à $\dfrac{k_x}{n_e^2} E_x + \dfrac{k_z}{n_o^2} E_z = 0$ montre que $\overrightarrow{E}_e$ est tangent à l'ellipsoïde (car orthogonal à sa normale) et donc que $\overrightarrow{\mathcal{P}}_e$, qui indique la direction de propagation de l'énergie, n'est pas parallèle à $\vec{k}$.

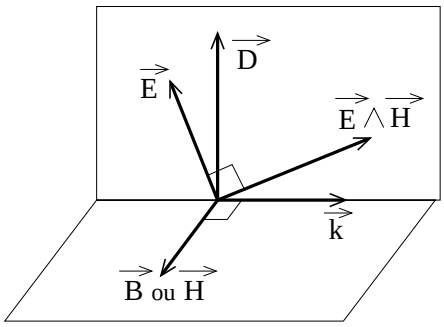


FIGURE 13

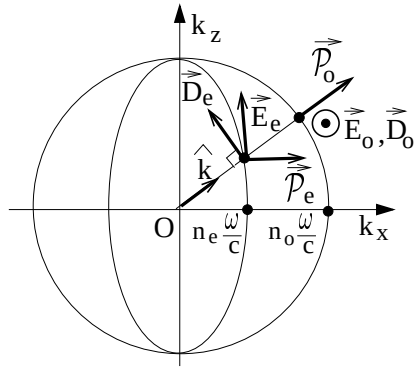


FIGURE 14

▷ Choix des solutions onde plane

Une relation de dispersion peut être satisfaite d'une infinité de façons. C'est la physique qui détermine le choix. Considérons l'exemple $\omega = -iDk^2$ de la **diffusion** à une dimension. Si k est réel, alors

$$\Re e \left[A\, e^{-i(\omega t - kx)} \right] = a\, e^{-Dk^2 t} \, \cos(kx - \varphi)$$

décrit l'amortissement "sur place" d'une perturbation sinusoïdale (spatialement) présente à $t = 0$ dans le milieu. Si ω est réel, $k = \left(\dfrac{i\omega}{D} \right)^{\frac{1}{2}} = \pm \dfrac{1+i}{\delta}$ (avec $\delta = \left(\dfrac{2D}{\omega} \right)^{\frac{1}{2}}$) ; la solution

$$\Re e \left[A\, e^{-i(\omega t - kx)} \right] = a\, e^{\mp \frac{x}{\delta}} \, \cos\Big(\omega t \mp \frac{x}{\delta} + \varphi \Big)$$

correspond alors à une perturbation de pulsation ω introduite à l'entrée d'un milieu semi-

infini ($[0,\infty]$ pour le signe "$-$" ou $[-\infty,0]$ pour le signe "$+$"), et qui, en "progressant" dans le milieu, s'amortit sur une distance δ (longueur de pénétration en thermique, épaisseur de peau en électromagnétisme, *etc.*).

De manière générale ω est réel si l'onde est générée par une source monochromatique (extérieure à la région considérée). Notons que dans ce cas, même si la relation de dispersion est réelle, $\vec{k}$ peut être complexe. C'est le cas lors de la **réflexion totale** d'une onde à la surface de séparation $z=0$ de deux milieux ($i=1,2$) isotropes pour lesquels $\omega^2 = v_i^2 k^2$ avec $v_1 < v_2$ (figure 15) ; pour une onde arrivant sous une incidence θ_1 telle que $\sin\theta_1 > \dfrac{v_1}{v_2}$, la loi de Descartes impose $k_{2x} = k_{1x} = \dfrac{\omega}{v_1}\sin\theta_1$ et alors $k_{2z}^2 = \dfrac{\omega^2}{v_2^2} - k_{2x}^2 < 0$. L'onde dans le milieu 2 ($z<0$) est une onde évanescente qui se propage parallèlement au plan $z=0$ et dont l'amplitude décroît comme $\exp|k_{2z}|z$.

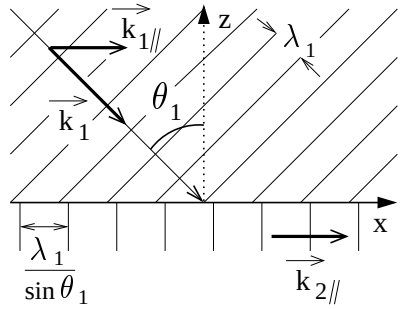


FIGURE 15

▷ Solution générale

Toute solution des E.D.P. linéaires à coefficients constants rencontrées en physique, qu'elles concernent un milieu infini ou limité (avec des C.L.), est une **somme (continue ou discrète) d'ondes planes** satisfaisant la relation de dispersion associée à l'E.D.P. considérée. On décrit ci-dessous quelques exemples de combinaisons d'ondes planes souvent rencontrées en physique.

EXEMPLE 1. **Paquet d'ondes**. Pour voir ce qui se rapproche le plus d'une solution onde plane, sommons des ondes dont les vecteurs d'onde $\vec{k} = \vec{k}_0 + \overrightarrow{K}$ sont proches de $\vec{k}_0$:

$$\Psi(\vec{r},t) = \iiint a(\vec{k}_0+\overrightarrow{K})\, e^{-i\left(\omega(\vec{k}_0+\overrightarrow{K})t - (\vec{k}_0+\overrightarrow{K})\cdot\vec{r} + \varphi(\vec{k}_0+\overrightarrow{K})\right)}\, \mathrm{d}^3K \quad (|\overrightarrow{K}| \simeq 0)) \ ,$$

avec $\omega(\vec{k})$ donnée par la relation de dispersion de l'E.D.P. considérée. En posant $\omega(\vec{k}_0) = \omega_0$ et $\varphi(\vec{k}_0) = \varphi_0$, et en effectuant les développements au premier ordre $\omega(\vec{k}_0+\overrightarrow{K}) = \omega_0 + \vec{v}_g \cdot \overrightarrow{K}$ et $\varphi(\vec{k}_0+\overrightarrow{K}) = \varphi_0 + \vec{r}_0 \cdot \overrightarrow{K}$ (avec $\vec{v}_g = \left.\dfrac{\partial\omega(\vec{k})}{\partial\vec{k}}\right|_{\vec{k}=\vec{k}_0}$ et $\vec{r}_0 = \left.\dfrac{\partial\varphi(\vec{k})}{\partial\vec{k}}\right|_{\vec{k}=\vec{k}_0}$) on obtient :

$$\Psi(\vec{r},t) = e^{-i(\omega_0 t - \vec{k}_0\cdot\vec{r} + \varphi_0)}\, f(\vec{r}-\vec{v}_g t - \vec{r}_0) \text{ avec } f(\vec{r}) = \iiint a(\vec{k}_0+\overrightarrow{K})\, e^{i\overrightarrow{K}\cdot\vec{r}}\, \mathrm{d}^3K \ \ (|\overrightarrow{K}| \simeq 0).$$

Ψ représente une onde plane modulée en amplitude dont l'enveloppe f, lentement variable, se déplace à la **vitesse de groupe** $\vec{v}_g(\vec{k}_0)$. (La figure 16a représente un paquet à une dimension de largeur Δz reliée à la dispersion Δk par $\Delta z\,\Delta k \simeq 1$.)

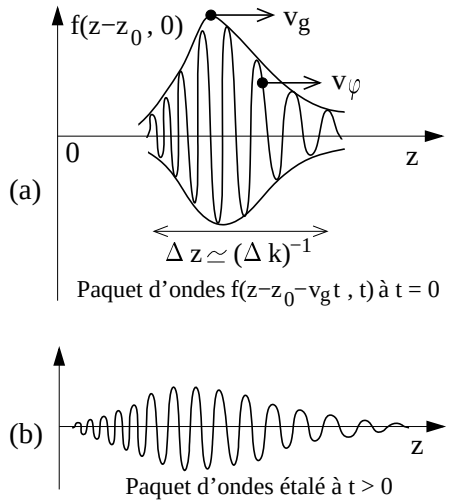


FIGURE 16

REMARQUES. 1) Dans un **milieu anisotrope** $\vec{v}_g$ égale à $\overrightarrow{\text{grad}}_{\vec{k}}\,\omega$, et donc perpendiculaire à la surface des vecteurs d'onde $\omega(\vec{k}) = \text{C}^{\text{ste}}$, n'est pas parallèle à $\vec{k}$; la figure 17 illustre la réfraction "extraordinaire" sous incidence normale ($\vec{k}_{\|} = 0$ continu) d'un paquet dans un tel milieu.

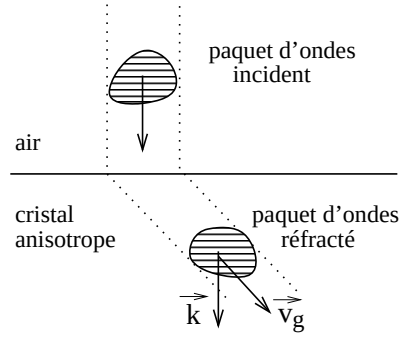


FIGURE 17

2) Dans un milieu isotrope (ou à une dimension) on a en général $v_g = \frac{\mathrm{d}\omega}{\mathrm{d}k} \neq v_\varphi = \frac{\omega}{k}$; par exemple en optique ($k = n(\omega)\,\frac{\omega}{c}$), pour des vagues en haute mer ($\omega^2 = kg$) et pour des particules libres quantiques ($\omega = \hbar\,\frac{k^2}{2m}$), on obtient respectivement $v_g^{-1} = v_\varphi^{-1} + \frac{\omega}{c}\,\frac{\mathrm{d}n}{\mathrm{d}\omega}$ (en pratique $\frac{v_\varphi - v_g}{v_\varphi} \simeq 10^{-1}$), $v_g = \frac{1}{2}\,v_\varphi$ et $v_g = 2\,v_\varphi$.

3) La dépendance de $\vec{v}_g$ vis-à-vis de $\vec{k}$ (à une dimension $\Delta v_g = \omega''(k)\,\Delta k$) est la cause de l'étalement ($\Delta v_g\,t$) des paquets et de la séparation (**dispersion**) des fréquences en son sein (figure 16b ; *cf.* aussi section 5.1.3).

4) La figure 18 montre l'arrivée à $t = 0$ en $\vec{r}_0 = 0$ d'un paquet d'ondes sur le plan $z = 0$ dans des conditions de réflexion totale pour les ondes qui le composent (coefficients de réflexion $r(\vec{k}) = e^{-i\Phi(\vec{k})}$). La formule donnant le paquet réfléchi diffère de celle du paquet incident par les changements $\vec{k}(k_x, k_y, k_z) \to \vec{k}'(k_x, k_y, -k_z)$ et $\varphi \to \varphi + \Phi$. Sa position (fictive) à $t = 0$ est donc $\vec{r}_0\,' = \overrightarrow{\text{grad}}_{\vec{k}'}\,\Phi$; ceci explique que le paquet réfléchi est décalé en position (**effet Goos Hänchen**) et en temps par rapport à ce que donnerait une réflexion "géométrique instantanée".

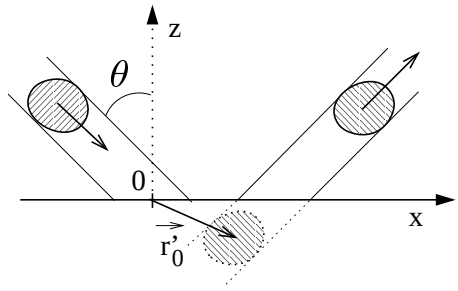


FIGURE 18

5) Signalons enfin un résultat très général et non trivial qu'on peut, en considérant la somme de deux ondes avec ω et $\vec{k}$ voisins, déduire de la L.C. de l'énergie $\partial_t e + \text{div}\vec{j} = 0$ et du caractère quadratique en ψ de e et $\vec{j}$: $\mathrm{d}\omega\ < e > -\mathrm{d}\vec{k}\cdot < \vec{j} >= 0$. Ce résultat entraîne que la vitesse de groupe $\vec{v}_g(\vec{k}_0)$ est égale au rapport des moyennes temporelles $\frac{<\vec{j}>}{<e>}$ calculé pour l'onde plane $(\omega_0, \vec{k}_0)$ (**vitesse de l'énergie**). De l'inégalité $< e > \geq | < \vec{j} > | c^{-1}$ (équivalente à $< e >> 0$ dans tous les référentiels), on déduit $|\vec{v}_g| \leq c$.

EXEMPLE 2. **Ondes guidées**. Si on fixe ω et $k_z = k\cos\theta$, et qu'on somme sur k_x et k_y en respectant la relation de dispersion (par exemple $\omega^2 = v^2\vec{k}^2$), on obtient des solutions du type $f(x, y)\, e^{-i(\omega t - k_z z)}$. Elles concernent des systèmes (**guides d'ondes**) invariants par translation selon z. Un exemple simple est la somme de deux ondes $A_\pm\, e^{-i(\omega t \mp k_x x - k_z z)}$ se propageant dans une région $x \in [0, a]$ (figure 19).

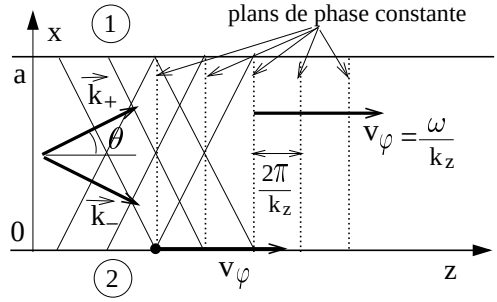


FIGURE 19

Les C.L. sont $A_+ = r_1\, A_-$ en $x = 0$ et $A_-\, e^{-ik_x a} = r_2\, A_+\, e^{ik_x a}$ en $x = a$, avec $r_{1,2} = e^{-i\varphi_{1,2}}$ coefficients de réflexion totale sur ces plans ; $\varphi_{1,2}$ dépendent en général de k_x et k_z et des milieux 1 et 2. Elles conduisent à la relation de quantification $2k_x a = \varphi_1 + \varphi_2 \overset{\text{def}}{=} \varphi$ (modulo 2π) pour k_x, et à la relation de dispersion $\omega^2 = v^2\left(k_x^2(k_z) + k_z^2\right)$.

(Par exemple $k_x = n\,\dfrac{\pi}{a}$ si $r_{1,2} = -1$.) Pour ces ondes : $v_\varphi = \dfrac{\omega}{k_z} = \dfrac{v}{\cos\theta}$ et (petit calcul) $v_g = \dfrac{\mathrm{d}\omega}{\mathrm{d}k_z} = v\,\cos\theta + v\,\sin\theta\,\dfrac{\partial\varphi}{\partial k_z}\left(2a - \dfrac{\partial\varphi}{\partial k_x}\right)^{-1}$; $v_g = v\,\cos\theta$ si $r_{1,2} = -1$.

Remarque. Si on somme des ondes planes de vecteurs d'onde ($k_x = k_\perp\,\cos\varphi_{\vec{k}}$, $k_y = k_\perp\,\sin\varphi_{\vec{k}}$, $k_z = k\,\cos\theta$), avec ω, k, θ fixés et $\varphi_{\vec{k}}$ variable de sommation, chaque onde étant affectée d'un poids $(2\pi)^{-1}\,e^{-im\varphi_{\vec{k}}}$, on obtient en coordonnées polaires :

$$f(r,\varphi) = (2\pi)^{-1}\int_0^{2\pi} e^{-im\varphi_{\vec{k}}}\,e^{ik_\perp r\,\cos(\varphi-\varphi_{\vec{k}})}\,\mathrm{d}\varphi_{\vec{k}} = e^{-im\varphi}\,J_m(k_\perp r)\ ;$$

(les **fonctions de Bessel** sont définies par $J_m(u) = (2\pi)^{-1}\int_0^{2\pi} e^{-im\varphi}\,e^{iu\,\cos\varphi}\,\mathrm{d}\varphi$.) Ces fonctions f qui vérifient $(\partial_x^2 + \partial_y^2 + k_\perp^2)\,f = 0$ (comme les fonctions $e^{i(k_x x + k_y y)}$) interviennent naturellement dans la physique des guides d'ondes cylindriques.

EXEMPLE 3. **Ondes sphériques** (figure 20). Une somme d'ondes planes monochromatiques de même amplitude et dont les vecteurs $\vec{k}$ ($|\vec{k}|$ fixé) sont proches de l'axe z ((k_x, k_y) dans un "petit" domaine $\mathcal{D}$ autour de $(0,0)$) s'écrit :

$$\Psi(\vec{r},t) = \iint_{\mathcal{D}} e^{-i(\omega t - k_x x - k_y y - k_z z)}\,\mathrm{d}k_x\,\mathrm{d}k_y \quad \text{avec} \quad k_z \simeq k - \frac{k_x^2 + k_y^2}{2k}\ .$$

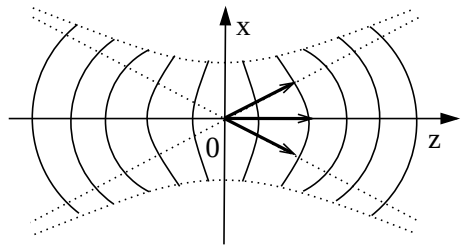


FIGURE 20

En utilisant l'approximation de la phase stationnaire (*cf.* section 5.1.3), le lecteur montrera que, pour z "grand", on a $\Psi \simeq 0$ si $\left(k_x^* = \dfrac{x}{z}k, k_y^* = \dfrac{y}{z}k\right) \notin \mathcal{D}$, et

$$\Psi \propto z^{-1}\,\exp -i\Big(\omega t - k\big(z + \frac{x^2+y^2}{2z}\big)\Big)$$

sinon. Pour $z > 0$ (resp. $z < 0$) il s'agit d'une onde sphérique divergente (resp. convergente) en O ; le passage par O entraîne un **déphasage de π** (lié au signe de z^{-1}).

8.2.2 Equations de diffusion et de propagation ; fonctions de Green ; ondes stationnaires

Pour résoudre ces équations, on utilise la dépendance linéaire de la solution vis-à-vis des C.I. (prises à $t = 0$) lorsqu'il n'y a pas de source. On montre ensuite que la solution causale en présence de source se déduit facilement du cas sans source. Cette démarche est celle suivie pour les E.D. à la section 6.2.1, la notion de fonction de Green généralisant aux E.D.P. celle de réponse impulsionnelle.

▷ E.D.P. de diffusion : $\partial_t f(x,t) - D\,\partial_x^2 f(x,t) = g(x,t)$ (à une dimension)

Si $g = 0$ on vérifie facilement que l'E.D.P. admet, pour $x \in [-\infty, \infty]$ et $t > 0$, la solution particulière :

$$G(x,t) = (4\pi D t)^{-\frac{1}{2}} \exp -\frac{x^2}{4Dt} \ .$$

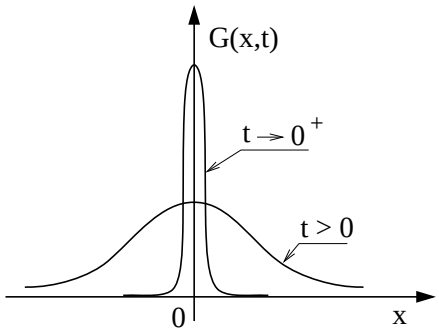


FIGURE 21

Il s'agit d'une fonction gaussienne centrée en 0, normalisée et d'écart type $\sqrt{2Dt}$ qui s'étale sans se propager quand t croît (figure 21). Quand $t \to 0^+$ elle est au contraire de plus en plus piquée et tend vers $\delta(x)$. On en déduit que la fonction

$$f(x,t) = \int_{-\infty}^{\infty} G(x - x', t)\, \varphi_0(x')\, \mathrm{d}x'$$

est, pour $t > 0$, la solution satisfaisant la C.I. $f(x,0) = \varphi_0(x)$.

Si $g \neq 0$ (**diffusion avec source**) la solution causale ("dont les variations ne précèdent pas celles de la source") est alors :

$$f(x,t) = \int_{-\infty}^{t} \left[\int_{-\infty}^{\infty} G(x - x', t - t')\, g(x', t')\, \mathrm{d}x' \right] \mathrm{d}t' \ .$$

DÉMONSTRATION : comme la fonction entre crochets satisfait (comme G) $\partial_t[\cdots] = D\,\partial_x^2[\cdots]$, et vaut $g(x,t)$ pour $t = t'$, on obtient $\partial_t f = g + \int_{-\infty}^{t} \partial_t[\cdots]\,\mathrm{d}t' = g + D\,\partial_x^2 f$. Cas particulier : si $g(x,t) = \delta(t)\,\delta(x)$, alors $f(x,t) = H(t)\,G(x,t)$; cette réponse à une impulsion en $x = 0$ à $t = 0$ est appelée **fonction de Green** du problème de diffusion à une dimension.

REMARQUE. A n dimensions les résultats précédents restent valables ; il suffit de remplacer x par $\underline{x}$ et $(4\pi D t)^{-\frac{1}{2}}$ par $(4\pi D t)^{-\frac{n}{2}}$. Pour l'**E.D.P. de Schrödinger** on remplace D par $\dfrac{i\hbar}{2m}$.

▷ E.D.P. de propagation à une dimension : $\partial_t^2 f(x.t) - v^2\,\partial_x^2 f(x,t) = g(x,t)$

Pour $g = 0$ la solution générale $f(x,t) = f_1(x - vt) + f_2(x + vt)$, somme d'ondes propagatives vue à la section 7.1.4, s'écrit aussi en fonction des C.I. $f(x,0) = \varphi_0(x)$ et $\partial_t f(x,t)|_{t=0} = \varphi_1(x)$ (vérification immédiate) :

$$f(x,t) = \frac{1}{2}\left(\varphi_0(x - vt) + \varphi_0(x + vt)\right) + \frac{1}{2v}\int_{x-vt}^{x+vt} \varphi_1(x')\,\mathrm{d}x' \ .$$

Pour une corde vibrante le premier terme décrit l'évolution d'une "perturbation de position" (obtenue en écartant la corde de sa position d'équilibre puis en la lâchant ; figure 22a), et le second (produit par t de la moyenne de $\varphi_1(x)$ sur l'intervalle $[x - vt, x + vt]$) celle d'une "perturbation de vitesse" (obtenue en soumettant la corde à une percussion ; figure 22b).

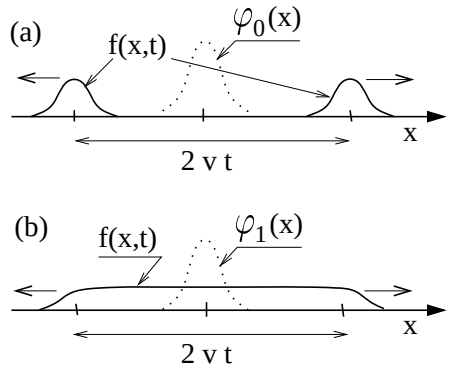


FIGURE 22

Si $g \neq 0$, il résulte de la relation entre f et φ_1 que la solution causale est :

$$f(x,t) = \int_{-\infty}^{t} \left[\frac{1}{2v}\int_{x-v(t-t')}^{x+v(t-t')} g(x',t')\,\mathrm{d}x'\right]\mathrm{d}t' \ .$$

DÉMONSTRATION : les propriétés de la fonction entre crochets, $\partial_t^2[\cdots] = v^2\,\partial_x^2[\cdots]$ ainsi que $[\cdots] = 0$ et $\partial_t[\cdots] = g(x,t)$ pour $t = t'$, entraînent $\partial_t f = \int_{-\infty}^{t} \partial_t[\cdots]\,\mathrm{d}t'$ puis $\partial_t^2 f = g + v^2\,\partial_x^2 f$. On notera que la solution s'écrit comme celle de la diffusion, la **fonction de Green** étant la porte $\dfrac{1}{2v}\,\Pi\!\left(\dfrac{x}{2vt}\right)$ au lieu de la gaussienne $G(x,t)$.

▷ E.D.P. de propagation à trois dimensions : $\partial_t^2 f(\vec{r},t) - v^2\,\Delta f(\vec{r},t) = g(\vec{r},t)$

Tous les résultats découlent du lemme suivant : soit $\varphi(\vec{r})$ une fonction arbitraire de $\vec{r}$; alors la fonction $\Phi(\vec{r},t) = t < \varphi >_{S(\vec{r},vt)}$, produit par $t > 0$ de la moyenne de φ sur la sphère centrée en $\vec{r}$ et de rayon vt, est telle que :

$$\partial_t^2\Phi(\vec{r},t) - v^2\,\Delta\Phi(\vec{r},t) = 0 \quad ; \quad \Phi(\vec{r},0) = 0 \quad ; \quad \partial_t\Phi(\vec{r},t)|_{t=0} = \varphi(\vec{r}) \quad ; \quad \partial_t^2\Phi(\vec{r},t)|_{t=0} = 0\,.$$

DÉMONSTRATION : la relation $\varphi \to \Phi$ étant linéaire on peut, en vertu de l'analyse de Fourier, se limiter au cas où φ est une onde plane $e^{i\vec{k}\cdot\vec{r}}$. Un point arbitraire sur la sphère $S(\vec{r}, vt)$ étant repéré par $\vec{r} + vt\hat{u}$ on a ($\mathrm{d}\Omega_{\hat{u}}$ angle solide élémentaire "autour"de $\hat{u}$) :

$$\Phi(\vec{r}, t) \stackrel{\text{def}}{=} \frac{t}{4\pi} \iint e^{i\vec{k}\cdot(\vec{r}+vt\hat{u})}\, \mathrm{d}\Omega_{\hat{u}} = e^{i\vec{k}\cdot\vec{r}}\, \frac{\sin \omega t}{\omega} \quad (\text{avec } \omega = vk) \ .$$

Les propriétés annoncées deviennent alors évidentes.
- Première conséquence : si $g = 0$, la solution $f(\vec{r}, t)$ s'exprime en fonction des C.I. $f(\vec{r}, 0) = \varphi_0(\vec{r})$ et $\partial_t f(\vec{r}, t)|_{t=0} = \varphi_1(\vec{r})$ sous la forme (vérification immédiate) :

$$f(\vec{r}, t) = \partial_t \Big(t < \varphi_0 >_{S(\vec{r},vt)} \Big) + t < \varphi_1 >_{S(\vec{r},vt)} \ .$$

- Deuxième conséquence : pour $g \neq 0$ la solution causale (causalité illustrée sur la figure 23), déduite de la relation ci-dessus entre f et φ_1, s'écrit (démonstration comme à une dimension) :

$$f(\vec{r}, t) = \int_{-\infty}^{t} \left[(t - t') \ < g(\vec{r}', t') >_{S\left(\vec{r}, v(t-t')\right)} \right] \mathrm{d}t' \ .$$

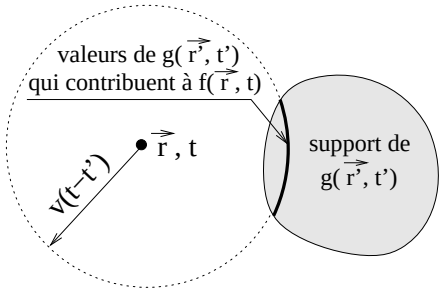


FIGURE 23

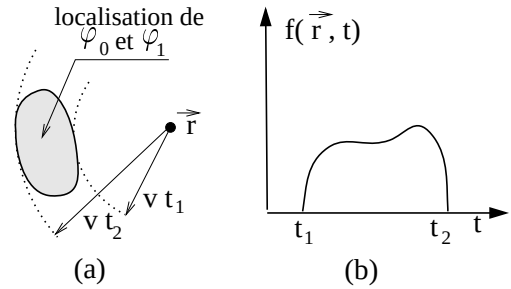


FIGURE 24

En posant $\vec{r}' = \vec{r} + R\hat{u}$ avec $R = v(t - t')$, elle devient

$$f(\vec{r}, t) = \int_0^{\infty} \left[\frac{1}{4\pi} \iint g\Big(\vec{r} + R\hat{u}, t - \frac{R}{v}\Big)\, \mathrm{d}\Omega_{\hat{u}} \right] \frac{R}{v}\, \frac{\mathrm{d}R}{v} \ ,$$

ou encore (forme plus habituelle de la **solution causale en physique**) :

$$f(\vec{r}, t) = \iiint \frac{g\big(\vec{r}', t - \frac{|\vec{r}-\vec{r}'|}{v}\big)}{4\pi v^2 |\vec{r} - \vec{r}'|}\, \mathrm{d}^3 r' \ .$$

(Par exemple la solution de $\Delta V - c^{-2}\, \partial_t^2 V = -\dfrac{\rho}{\epsilon_0}$ est écrite à la section 7.5.2). La **fonction de Green**, réponse à $\delta(\vec{r})\, \delta(t)$, est ici $H(t)\, (4\pi v^2 r)^{-1}\, \delta\Big(t - \dfrac{r}{v}\Big)$ ("pulse" sphérique naissant à $t = 0$ en $r = 0$).

REMARQUES (pour $g = 0$). 1) Il résulte de la première conséquence que, si φ_0 et φ_1 sont localisées (figure 24a), l'onde n'est présente en tout point $\vec{r}$ que pendant un intervalle de temps fini (figure 24b).

2) Dans le **cas statique** ($f(\vec{r},t) \equiv \varphi_0(\vec{r})$ et $\Delta\varphi_0 = 0$), elle entraîne aussi que la valeur $\varphi_0(\vec{r})$ est la moyenne de φ_0 sur toute sphère centrée en $\vec{r}$ (généralisation du résultat établi à la section 7.1.2).
3) Les résultats relatifs à la propagation à une (resp. deux) dimension se déduisent du cas à trois dimensions en considérant des fonctions φ_0 et φ_1 ne dépendant que de x (resp. x et y), et en intégrant toujours sur les trois variables. On vérifie par exemple que la moyenne sur la sphère $S(\vec{r},vt)$ devient la moyenne sur le segment $[x-vt, x+vt]$ à une dimension.

▷ Introduction de conditions aux limites (C.L.) ; ondes stationnaires

Les solutions précédentes supposent un espace infini. Si on impose des C.L. indépendantes du temps, les ondes stationnaires $f(\vec{r},t) = f(\vec{r})\,h(t)$ (**produit d'une fonction de l'espace par une fonction du temps**), solution des équations en absence de source ($g = 0$), jouent un rôle important.

Par exemple l'E.D.P. $(\partial_t^2 - v^2\partial_x^2)f(x)\,h(t) = 0$ conduit à $\dfrac{\ddot{h}}{h} = v^2\,\dfrac{f''}{f} = -\omega^2 = \mathrm{C^{ste}}$ (en divisant par fh). (0n suppose h borné pour $t \to \pm\infty$). Les variables t et x étant indépendantes on en déduit (si $\omega \neq 0$) :

$$f(x)\,h(t) = A\,\cos(kx+\varphi)\,\cos(\omega t+\phi) \quad \text{avec} \quad k = \frac{\omega}{v}\ .$$

Pour une corde vibrante attachée en $x = 0$ et $x = L$, les C.L. fixent la dépendance spatiale en $\sin k_n x$ avec $k_n = n\dfrac{\pi}{L}$ ($n > 0$). La solution générale est alors $f(x,t) = \sum_n \sin k_n x\,\big(a_n\,\cos\omega_n t + b_n\,\sin\omega_n t\big)$; les coefficients a_n (resp. b_n) sont déterminés par les développements en série de Fourier de $f(x,0)$ (resp. $\partial_t f(x,t)|_{t=0}$).

De même, pour l'E.D.P. de diffusion $(\partial_t - D\,\partial_x^2)\,f(x)\,h(t) = 0$, on obtient $\dfrac{\dot{h}}{h} = D\,\dfrac{f''}{f} = -Dk^2$ (h bornée à $t \to \infty$), d'où $f(x)\,h(t) = A\,\cos(kx+\varphi)\,e^{-Dk^2t}$ (si $k \neq 0$). Pour des C.L. nulles en $x = 0$ et $x = L$ la solution générale est $f(x,t) = \sum_n a_n\,\sin k_n x\,e^{-Dk_n^2 t}$ et les C.I. donnent $a_n = \dfrac{2}{L}\displaystyle\int_0^L f(x,0)\,\sin k_n x\,\mathrm{d}x$. (Pour des C.L. constantes non nulles il faudrait ajouter un terme $a_0 x + b_0$ correspondant à $k = 0$.)

Ce type de solutions s'étend à trois dimensions et à d'autres équations. Par exemple les E.D.P. de propagation en milieux linéaires inhomogènes $\partial_t^2 f = v^2(\vec{r})\,\Delta f$, ou l'E.D.P. de Schrödinger $i\hbar\,\partial_t f = -\dfrac{\hbar^2}{2m}\,\Delta f + V(\vec{r})\,f$, conduisent aux **E.D.P. spatiales**

$$\big(\Delta + k^2(\vec{r})\big)\,f(\vec{r}) = 0$$

avec respectivement $k(\vec{r}) = \dfrac{\omega}{v(\vec{r})}$ (pour $h(t) \propto \cos(\omega t+\varphi)$) et $k^2(\vec{r}) = \dfrac{2m}{\hbar^2}\,\big(E - V(\vec{r})\big)$ (pour $h(t) \propto \exp -i\dfrac{Et}{\hbar}$).

Fonction de Green avec C.L.. Considérons l'exemple de l'équation de diffusion à une dimension ci-dessus dans $[0,L]$ avec C.L. nulles. L'expression des a_n ci-dessus permet d'écrire la relation linéaire entre la solution $f(x,t)$ et la C.I. $f(x,0)$ sous la forme :

$$f(x,t) = \textstyle\int_0^L G(x,x',t)\,f(x',0)\,\mathrm{d}x' \quad \text{avec} \quad G(x,x',t) = \frac{2}{L}\sum_n \sin k_n x\,\sin k_n x'\,e^{-Dk_n^2 t}\ .$$

La fonction $G(x,x',t)$ qui vérifie $\partial_t G = D\,\partial_x^2 G$, les C.L., et qui est égale à $\delta(x-x')$ pour $t = 0$,

est la fonction de Green du problème de diffusion dans $[0, L]$. Elle conduit à la solution $f(x,t) = \int_{-\infty}^{t} \left[\int_0^L G(x,x',t-t')\, g(x',t')\, \mathrm{d}x'\right] \mathrm{d}t'$ du même problème en présence de source.

8.2.3 Six E.D.P. liées à une loi de conservation

On considère les bilans (sans source)

$$\partial_t f + \partial_x j = 0 \text{ (à une dimension) et } \partial_t f + \mathrm{div}\vec{j} = 0 \text{ (à trois dimensions)} .$$

Ces bilans dans lesquels f est en général une densité (linéaire ou volumique) de grandeur additive conservée, conduisent à une E.D.P. pour f lorsque l'intensité du courant j, ou sa densité volumique $\vec{j}$, est liée à f.

EXEMPLE 1. Si j est une fonction $j(f)$ on obtient l'**E.D.P. d'onde simple** non linéaire

$$\partial_t f + c(f)\, \partial_x f = 0 \quad (\text{avec } c(f) = j'(f))$$

qui peut conduire à une onde de choc (*cf.* exemple de la section 7.2.4).

EXEMPLE 2. Si $j = -D\, \partial_x f$ ou $\vec{j} = -D\, \overrightarrow{\mathrm{grad}} f$ on obtient l'équation de diffusion. Un exemple est l'**équation de la "chaleur"**

$$\rho c\, \partial_t T = K\, \Delta T ,$$

qui correspond à un bilan d'énergie interne $\partial_t(\rho e) + \mathrm{div}\vec{j}_{\mathrm{th}} = 0$ pour un solide, avec $e = \int^T c(T')\, \mathrm{d}T'$ et $\vec{j}_{\mathrm{th}} = -K\, \overrightarrow{\mathrm{grad}} T$.

Remarque. Le lecteur établira et interprètera le bilan relatif à l'entropie massique $s = \int^T T'^{-1} c(T')\, \mathrm{d}T'$:

$$\partial_t(\rho s) + \mathrm{div}\vec{j}_{\mathrm{S}} = \tau_{\mathrm{S}} \quad \text{avec} \quad \vec{j}_{\mathrm{S}} = T^{-1}\vec{j}_{\mathrm{th}} \quad \text{et} \quad \tau_{\mathrm{S}} = KT^{-2}(\overrightarrow{\mathrm{grad}} T)^2 .$$

EXEMPLE 3. Si $j(x,t) = V(x,t)\, f(x,t)$ l'E.D.P. linéaire inhomogène

$$\partial_t f + \partial_x(Vf) = 0$$

décrit l'**entraînement d'une grandeur** en tout point x à l'instant t par un écoulement dont le champ de vitesse est $V(x,t)$. En effet, en suivant le mouvement défini par $\dot{x} = V(x,t)$, on a $\mathrm{d}_t\big(f(x,t)\,\mathrm{d}x\big) = (\mathrm{d}_t f + f\, \partial_x V)\,\mathrm{d}x = \big(\partial_t f + \partial_x(fV)\big)\,\mathrm{d}x = 0$ (*cf.* sections 7.1.3 et 7.2.3).

Remarque. La même notion d'entraînement s'applique à la L.C. $\partial_t f + \partial_{\underline{x}}\,(\underline{V}(\underline{x},t)\, f) = 0$ en dimension arbitraire. Si de plus $\underline{V}$ est à divergence nulle ($\partial_{\underline{x}} \cdot \underline{V} = 0$), alors $\mathrm{d}_t f = 0$; un exemple est donné par l'**E.D.P. de Liouville** satisfaite par toute densité de probabilité $f(\underline{q},\underline{p},t)$ lorsque les points de l'espace de phase obéissent à une dynamique hamiltonienne : $\underline{V} \equiv (\partial_{\underline{p}} H, -\partial_{\underline{q}} H)$ (*cf.* section 9.1).

EXEMPLE 4. Citons également deux E.D.P. associant l'effet d'entraînement à d'autres effets. L'**E.D.P. de Fokker-Planck** introduite à la section 10.3.1 combine entraînement et diffusion. L'**E.D.P. de Boltzmann**

$$\partial_t f + \tfrac{\vec{F}}{m} \cdot \partial_{\vec{v}} f + \vec{v} \cdot \overrightarrow{\nabla} f = -\tfrac{f - f_0}{\tau} ,$$

relative à une densité de probabilité dans l'espace de phase $f(\vec{r},\vec{v},t)$ à une seule particule, combine entraînement (mouvement d'un point $(\vec{r},\vec{v})$ sous l'action d'une force extérieure : $\dot{\vec{r}} = \vec{v}$ et $\dot{\vec{v}} = \tfrac{\vec{F}}{m}$) et relaxation vers une distribution d'équilibre (en général du type $f_0(\vec{v})$, de Maxwell ou Fermi-Dirac). Sous sa forme linéarisée, elle sert à décrire de nombreux **phénomènes de transport** ; par exemple si $\vec{F} = q\vec{E}$ est constante, la conductivité électrique γ, définie par $\vec{j} = \gamma\vec{E}$, s'obtient en écrivant :

$$f = f_0 - \tau\, \tfrac{q}{m}\, \vec{E} \cdot \partial_{\vec{v}} f_0 = f_0 + \delta f \quad \text{et} \quad \vec{j} = \iiint q\, \vec{v}\, \delta f\, \mathrm{d}^3 v .$$

8.2.4 Trois exemples d'E.D.P. non linéaires ; ondes solitaires

Ces E.D.P. se rencontrent dans tous les domaines (hydrodynamique, optique, plasmas, chimie, biologie...) avec des applications dans la vie courante (écoulements d'eau, houle, organisation des nuages, propagation des incendies...). En général les nonlinéarités sont dues aux propriétés des milieux et, en hydrodynamique, à la présence de la dérivée particulaire $\mathrm{d}_t = \partial_t + \vec{v} \cdot \overrightarrow{\nabla}$. Si elles sont petites on peut montrer, comme pour les oscillateurs (*cf.* section 6.5.1), par une méthode de perturbation, que leurs effets sont la dépendance $\omega(k, a)$ de la relation de dispersion vis-à-vis de l'amplitude a et la présence d'harmoniques.

On considère ci-dessous trois exemples "classiques" d'E.D.P. non linéaires à une dimension et leurs solutions particulières du type $f(u = x - Vt)$. Ces ondes qui se propagent sans déformation à vitesse constante V, et sont constantes (ou nulles) pour u infini, sont appelées **ondes solitaires**.

▷ E.D.P. de Sine Gordon $\partial_t^2 f = -\omega_0^2 \sin f + v^2 \partial_x^2 f$

Elle est la limite continue du système d'équations

$$ml^2 \ddot{\theta}_n = -mgl \sin\theta_n - C(\theta_n - \theta_{n-1}) - C(\theta_n - \theta_{n+1})$$

décrivant des pendules simples oscillant dans un plan perpendiculaire à l'axe horizontal "x" et couplés linéairement par des ressorts de torsion portés par cet axe. Dans la suite on fera $\omega_0 = v = 1$, le lecteur pouvant rétablir dimensionnellement ω_0 et v dans toutes les formules (*cf.* chapitre 1). L'équation pour $f(u)$, qui s'écrit $(1 - V^2)\, f'' = \sin f$, est aussi celle d'un mobile de masse $1 - V^2$ se déplaçant dans le potentiel $U(f) = 1 + \cos f$.

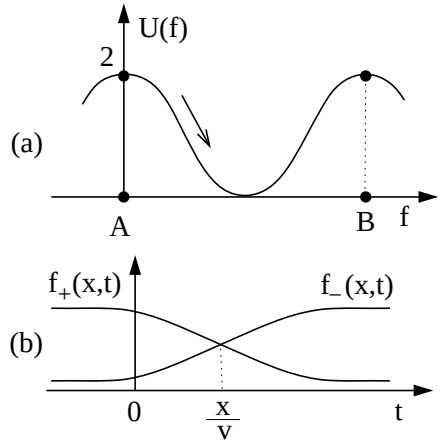


FIGURE 25

On voit alors sur la figure 25a que f ne tend vers une constante à l'infini que si ce mobile va de A à B ou de B à A (modulo 2π). L'intégrale première d'"énergie" $(1-V^2)\,\frac{f'^2}{2} + U(f) = 2$ conduit à $(1 - V^2)^{\frac{1}{2}}\, \mathrm{d}f = \pm 2 \sin\frac{f}{2}\, \mathrm{d}u$ et aux solutions (figure 25b) :

$$f_\pm(x - Vt) = 4\,\mathrm{Arctg}\left[\exp \pm \frac{x - Vt}{\sqrt{1 - V^2}}\right] .$$

En pratique on obtient ces ondes en faisant faire un tour complet "à la main" au pendule situé à une extrémité d'une longue chaîne.

REMARQUE. Si on fait faire un tour complet aux deux bouts de la chaîne on observe que les deux ondes, qui se propagent en sens inverses, reprennent asymptotiquement leur forme initiale après s'être "traversées" ; on les nomme "solitons". Un exemple de solution superposition non linéaire de deux solitons $f_+(x \pm Vt)$ est (vérification laissée au lecteur) : $f_2(x,t) = 4\,\mathrm{Arctg}\Big[V\,\sinh\frac{x}{\sqrt{1-V^2}}\,\cosh^{-1}\frac{Vt}{\sqrt{1-V^2}}\Big] + 2\pi$.

▷ E.D.P. de Korteweg-De Vries $\partial_t f + (c_0 + c_1 f)\,\partial_x f + \nu\,\partial_x^3 f = 0$

Elle décrit des ondes de surface dans un canal peu profond ($c_0 = (gh)^{\frac{1}{2}}$; *cf.* section 8.1.2). Dans la suite on pose $c_0 = c_1 = \nu = 1$. En procédant par intégrations successives, et en supposant que $f \to 0$ pour $u \to \pm\infty$, on obtient l'équation $\frac{1}{2}\,f'^2 + U(f) = 0$ avec $U(f) = \frac{1-V}{2}\,f^2 + \frac{1}{6}\,f^3$, dont on déduit (petit calcul) :

$$f(x - Vt) = a\,\cosh^{-2}\Big[\sqrt{\tfrac{a}{12}}\,(x - Vt)\Big] \qquad \text{avec} \quad V = 1 + \tfrac{a}{3}\ .$$

La non déformation de cette onde solitaire tient au fait qu'elle est un compromis entre d'une part les paquets d'ondes solutions de l'E.D.P. linéarisée qui s'étalent (relation de dispersion $\omega = k - k^3$), et d'autre part l'onde solution de $\partial_t f + (1 + f)\,\partial_x f = 0$, équation qui conduit à une onde de choc (*cf.* section 7.1.1).

▷ E.D.P. de Burgers $\partial_t f + (c_0 + c_1 f)\,\partial_x f = D\,\partial_x^2 f$

Elle ajoute des effets de diffusion à l'équation d'onde simple. La solution onde solitaire (avec $c_0 = c_1 = 1$)

$$f(x - Vt) = a\Big(1 + \exp\tfrac{x - Vt}{l}\Big)^{-1} \qquad \text{avec} \quad V = 1 + \tfrac{a}{2} \quad \text{et} \quad l = 2\tfrac{D}{a}$$

montre que la diffusion a pour effet d'étaler l'onde de choc.

8.3 E.D.P. « SPATIALES » IMPLIQUANT L'OPÉRATEUR LAPLACIEN

8.3.1 Exemples et analogies physiques ; conditions aux limites

▷ E.D.P. de Laplace $\Delta f = 0$

Elle concerne des fonctions "potentiel" ($f \equiv V, \varphi, n, T, \dots$) et apparaît chaque fois qu'un champ a un rotationnel et une divergence nuls, par exemple en **électrostatique** ($\overrightarrow{E} = -\overrightarrow{\mathrm{grad}}\,V$ et $\mathrm{div}\,\overrightarrow{E} = 0$), en **électrocinétique** (idem avec $\vec{j} = \gamma\overrightarrow{E}$), en **gravitation** ($\vec{g} = -\overrightarrow{\mathrm{grad}}\,V$), en **hydrodynamique** des écoulements irrotationnels incompressibles ($\vec{v} = \overrightarrow{\mathrm{grad}}\,\varphi$ et $\mathrm{div}\vec{v} = 0$), dans des problèmes stationnaires de **diffusion** ($\vec{j}_n = -D\,\overrightarrow{\mathrm{grad}}\,n$ et $\mathrm{div}\vec{j}_n = 0$), ou de **conduction thermique** ($\vec{j}_{\mathrm{th}} = -K\,\overrightarrow{\mathrm{grad}}\,T$ et $\mathrm{div}\vec{j}_{\mathrm{th}} = 0$). En **magnétostatique** ($\overrightarrow{\mathrm{rot}}\,\overrightarrow{B} = 0$ et $\mathrm{div}\overrightarrow{B} = 0$) on a $\Delta V_{\mathrm{mag}} = 0$ pour le potentiel scalaire magnétique défini (localement) par $\overrightarrow{B} = -\overrightarrow{\mathrm{grad}}\,V_{\mathrm{mag}}$, et $\Delta\overrightarrow{A} = 0$ pour le potentiel vecteur défini par $\overrightarrow{B} = \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}$ et $\mathrm{div}\overrightarrow{A} = 0$.

Cette identité mathématique entre différents domaines de la physique permet des analogies. Ainsi pour un tube de champ limité par deux équipotentielles, les quantités $C^{-1} = \dfrac{V_1 - V_2}{Q}$ (C capacité d'un condensateur), $R = \dfrac{V_1 - V_2}{I}$ (résistance électrique)

et $R_{\text{th}} = \dfrac{T_1 - T_2}{I_{\text{th}}}$ (résistance thermique) font appel à un même calcul, celui du rapport $\mathcal{R}$ de la circulation du champ (respectivement $\overrightarrow{E}$, $\vec{j}$ et $\vec{j}_{\text{th}}$) entre les équipotentielles et de son flux dans le tube : $\mathcal{R} = \epsilon_0 C^{-1}$, γR et $K R_{\text{th}}$. $\mathcal{R}$ ne dépend que de la géométrie et est homogène à $[L]^{-1}$. Pour des géométries planes (figure 26a), sphériques, cylindriques (deux cas différents correspondant aux figures 26b et 26c), on a respectivement : $V \sim x$ et $\mathcal{R} = \Delta x\, S^{-1}$ (S section du tube) ; $V \sim r^{-1}$ et $\mathcal{R} = \dfrac{1}{4\pi}\left(\dfrac{1}{r_1} - \dfrac{1}{r_2}\right)$; enfin $V \sim \ln r$ et $\mathcal{R} = (l\,\Delta\theta)^{-1} \ln \dfrac{r_2}{r_1}$ (l longueur du cylindre) ou $V \sim \theta$ et $\mathcal{R} = (\Delta\theta)\left(l \ln \dfrac{r_2}{r_1}\right)^{-1}$.

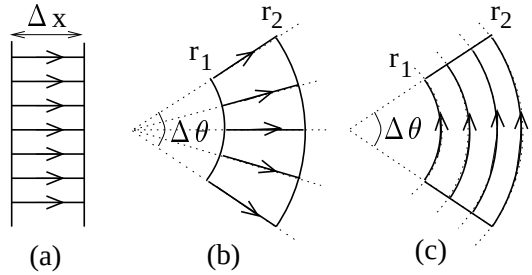


FIGURE 26

▷ Autres équations

L'**E.D.P. de Poisson** $\Delta f = g$ apparait dans les exemples précédents lorsque des sources statiques sont présentes, par exemple $\Delta V = -\dfrac{\rho}{\epsilon_0}$ ou $\Delta \overrightarrow{A} = -\mu_0 \vec{j}$. L'équation $\left(\Delta + k^2(\vec{r})\right) f = 0$ (**E.D.P. de Helmholtz** si $k^2 = \mathrm{C}^{\text{ste}} > 0$) apparait quand on recherche des solutions onde stationnaire à une équation de propagation ou de Schrödinger (*cf.* section 8.2.2). Souvent $k^2(\vec{r}) = k_0^2 + \delta k^2(\vec{r})$ où la perturbation $\delta k^2(\vec{r})$ décrit des "inhomogénéités" (par exemple $k_0^2 = \dfrac{2mE}{\hbar^2}$ et $\delta k^2(\vec{r}) = -\dfrac{2m}{\hbar^2} V(\vec{r})$).

▷ Conditions aux limites

En général les domaines dans lesquels les équations précédentes s'appliquent sont limités par des surfaces. Les conditions aux limites sont dites **C.L. de Dirichlet** si c'est f qui est donnée sur la surface, et **C.L. de Neuman** si c'est sa dérivée normale $\hat{n} \cdot \overrightarrow{\text{grad}} f$. En électrostatique c'est souvent le potentiel V qui est fixé (et constant) sur la surface, tandis qu'en hydrodynamique c'est le "courant" $\hat{n} \cdot \overrightarrow{\text{grad}} \varphi$ (égal à 0 si la surface limite le fluide). Dans certains domaines les deux types de C.L. (et même des conditions mixtes faisant intervenir l'impédance à la surface) peuvent être rencontrées. Si le domaine n'est pas limité par des surfaces, il faut préciser le comportement à l'infini de f ; par exemple, en quantique, la fonction d'onde est exponentiellement décroissante pour des **états liés**, et du type $e^{i\vec{k}_0 \cdot \vec{r}} + A(\vec{k}, \vec{k}_0)\, \dfrac{e^{ik_0 r}}{r}$ pour des **états de diffusion** ($\vec{k}_0$ est le vecteur d'onde incident et $\vec{k} = |\vec{k}_0|\, \hat{r}$ indique la direction d'observation de l'onde diffusée).

8.3.2 Unicité des solutions ; identité de Green ; séparation des variables

▷ Unicité des solutions de $\Delta f = g$

Pour des C.L. de Dirichlet ou des C.L. de Neuman il y a unicité (éventuellement modulo une constante) de la solution de l'E.D.P. $\Delta f = g$ dans un volume V limité par une ou plusieurs surfaces S. Si de plus $g = 0$, la solution minimise l'intégrale

$$I = \frac{1}{2} \iiint_V |\overrightarrow{\nabla} f|^2 \, \mathrm{d}^3 r \ .$$

I peut représenter une énergie (électrostatique, cinétique...), une puissance dissipée ou encore, à deux dimensions, l'aire d'une surface $z = f(x, y)$ s'appuyant sur une courbe fermée (lame d'eau savonneuse). Ces résultats découlent de l'identité :

$$\oiint_S G \, \overrightarrow{\nabla} F \cdot \overrightarrow{\mathrm{d}S} = \iiint_V G \, \Delta F \, \mathrm{d}^3 r + \iiint_V \overrightarrow{\nabla} F \cdot \overrightarrow{\nabla} G \, \mathrm{d}^3 r \ .$$

DÉMONSTRATION de l'unicité : supposons qu'il existe deux solutions f_1 et f_2 satisfaisant les mêmes C.L. et posons $F = G = f_2 - f_1$. Alors sur S soit $G = 0$ (Dirichlet), soit $\overrightarrow{\nabla} F \cdot \overrightarrow{\mathrm{d}S} = \overrightarrow{\nabla} F \cdot \hat{n} \, \mathrm{d}S = 0$ (Neuman) ; l'intégrale de surface est donc nulle et, comme $\Delta F = 0$, on en déduit $|\overrightarrow{\nabla} F|^2 = 0$ soit $F = \mathrm{C^{ste}}$ ($= 0$ pour Dirichlet). Démonstration du minimum (pour $g = 0$) : on pose $F = f$ et $G = \delta f$ (Dirichlet) ou $F = \delta f$ et $G = f$ (Neuman). L'intégrale de surface est alors nulle et, comme $\Delta f = 0$, on a $\iiint_V \overrightarrow{\nabla} f \cdot \overrightarrow{\nabla} (\delta f) \, \mathrm{d}^3 r = \delta I = 0$; l'extremum est un minimum car $I > 0$ est bornée inférieurement.

▷ Méthode des images

Elle est basée sur l'unicité de la solution. Si par exemple en électrostatique on veut résoudre $\Delta V = -\dfrac{\rho}{\epsilon_0}$ dans une région limitée par une surface S sur laquelle une C.L. est donnée, cette méthode consiste à introduire une répartition ρ^* de charges situées de l'autre côté de S qui, avec ρ, réalisent la condition sur S puis à résoudre $\Delta V = -\dfrac{\rho + \rho^*}{\epsilon_0}$ (sans plus se préoccuper de la C.L.).

EXEMPLE 1 (figure 27a). Si une charge q_1 est à l'extérieur d'une sphère métallique S centrée en O, de rayon R au potentiel $V_S = 0$, on détermine une charge $-q_2$ et sa position de sorte que l'équation $\dfrac{q_1}{r_1} - \dfrac{q_2}{r_2} = 0$ soit celle de la sphère (*cf.* section 2.2.2). Le potentiel à l'extérieur de S est celui dû à q_1 et $-q_2$. Si $V_S \neq 0$, on ajoute en O une charge $Q = 4\pi\epsilon_0 R V_S$.

EXEMPLE 2. Si la sphère métallique au potentiel $V = 0$ est en présence d'un champ extérieur constant $\overrightarrow{E}_0$ (figure 27 b), on remarque qu'en superposant à $\overrightarrow{E}_0$ le champ d'un dipole de moment $\vec{p} \parallel \overrightarrow{E}_0$ placé en O, le potentiel en tout point est :

$$V(r, \theta) = -E_0 r \, \cos\theta + \tfrac{p \, \cos\theta}{4\pi\epsilon_0 r^2} \ .$$

L'équipotentielle $V = 0$ coïncide avec S si $\vec{p} = 4\pi\epsilon_0 R^3 \overrightarrow{E}_0$. $V(r, \theta)$ décrit alors toutes les propriétés (lignes de champ, répartition de charges et pression à la surface de la sphère) du système “sphère plus

champ $\overrightarrow{E}_0$" pour $r \geq R$. Si sur S (supposée maintenant non métallique) on impose une condition de Neuman du type $E_r = -\partial_r V = 0$, on voit qu'on doit prendre $\vec{p} = -2\pi\epsilon_0 R^3 \overrightarrow{E}_0$. Cette dernière C.L., non réaliste pour un champ électrostatique, le devient pour la vitesse $\vec{v}$ d'un écoulement autour d'un obstacle sphérique (figure 28), ou pour le champ magnétique $\overrightarrow{B}$ à la surface d'un supraconducteur sphérique à l'intérieur duquel $\overrightarrow{B} = 0$ en raison de l'effet Meissner (*cf.* section 9.4).

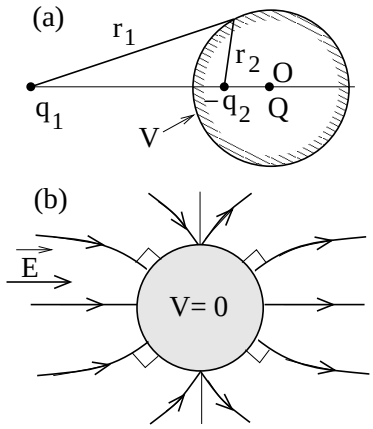


FIGURE 27

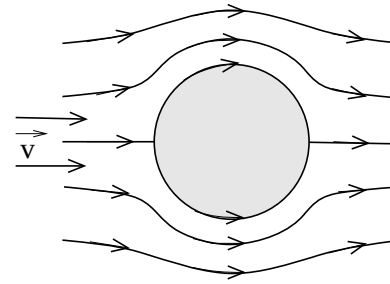


FIGURE 28

▷ Identité de Green et applications

Green Pour tout point $\vec{r}_0$ à l'intérieur d'un volume V limité par une surface S on a :

$$4\pi\, f(\vec{r}_0) = \oiint_S \left[\frac{e^{ik|\vec{r}-\vec{r}_0|}}{|\vec{r}-\vec{r}_0|} \overrightarrow{\nabla} f(\vec{r}) - \left(\overrightarrow{\nabla} \frac{e^{ik|\vec{r}-\vec{r}_0|}}{|\vec{r}-\vec{r}_0|} \right) f(\vec{r}) \right] \cdot \overrightarrow{\mathrm{d}S} - \iiint_V \frac{e^{ik|\vec{r}-\vec{r}_0|}}{|\vec{r}-\vec{r}_0|} \left(\Delta + k^2\right) f(\vec{r})\, \mathrm{d}^3 r \,.$$

DÉMONSTRATION (figure 29) : elle s'appuie sur l'expression intégrale de la formule de Green :

$$\oiint_S (F\, \overrightarrow{\nabla} G - G\, \overrightarrow{\nabla} F) \cdot \overrightarrow{\mathrm{d}S} = \iiint_V (F\, \Delta G - G\, \Delta F)\, \mathrm{d}^3 r \,.$$

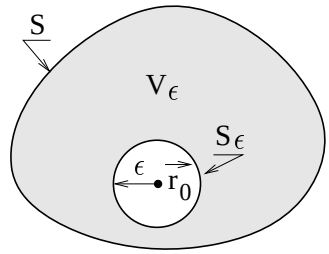


FIGURE 29

On pose alors $F = f$ et $G = |\vec{r} - \vec{r}_0|^{-1}\, e^{ik|\vec{r}-\vec{r}_0|}$, et on prend pour V le volume V_ϵ limité par S et la sphère S_ϵ (de rayon ϵ, centrée en $\vec{r}_0$ et orientée vers l'extérieur). Dans la limite $\epsilon \to 0$ l'intégrale sur S_ϵ de $G\, \overrightarrow{\nabla} F$ tend vers zéro ($G\, \overrightarrow{\nabla} F \sim \epsilon^{-1}$ et $S_\epsilon = 4\pi\epsilon^2$), celle de $F\, \overrightarrow{\nabla} G$ tend vers $4\pi\, f(\vec{r}_0)$ ($|\overrightarrow{\nabla} G| \sim \epsilon^{-2}$), et dans $V_\epsilon \to V$ on a $\Delta G = \frac{\mathrm{d}^2 G}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}G}{\mathrm{d}r} = -k^2 G$, d'où l'identité.

Citons quelques conséquences remarquables de cette identité.

1) Si $\Delta f = 0$, si $k = 0$ et si S est une sphère de rayon R centrée en $\vec{r}_0$, on retrouve facilement (car $|\vec{r}-\vec{r}_0|$ est constant sur S) le résultat de la section 8.2.2 : $f(\vec{r}_0) =< f >_{S(\vec{r}_0,R)}$.

2) Si $\Delta f = g$ (**E.D.P. de Poisson**), et dans l'hypothèse où l'intégrale de surface tend vers zéro lorsque S va à l'infini, elle donne :

$$f(\vec{r}) = -(4\pi)^{-1} \iiint |\vec{r}\,' - \vec{r}|^{-1}\, g(\vec{r}\,')\, \mathrm{d}^3 r' \ .$$

(On vérifie à posteriori que l'hypothèse sur l'intégrale de surface est justifiée lorsque g est localisée dans une région finie car f se comporte alors comme r^{-1} à l'infini.)

3) **Diffusion** : si f satisfait $(\Delta + k^2)\, f = g$ et se comporte comme une "onde sphérique sortante" à l'infini, on obtient :

$$f(\vec{r}) = -(4\pi)^{-1} \iiint |\vec{r}\,' - \vec{r}|^{-1}\, e^{ik|\vec{r}\,'-\vec{r}|}\, g(\vec{r}\,')\, \mathrm{d}^3 r' \ .$$

Une application intéressante concerne les problèmes de diffusion régis par l'équation

$$\left(\Delta + k_0^2 + \delta k^2(\vec{r})\right)\, f(\vec{r}) = 0$$

où $\delta k^2(\vec{r})$ peut être traité comme une perturbation. La solution non perturbée étant $f_0(\vec{r}) \propto e^{i\vec{k}_0\cdot\vec{r}}$, on pose $g(\vec{r}) = -\delta k^2(\vec{r})\, f(\vec{r}) \simeq -\delta k^2(\vec{r})\, e^{i\vec{k}_0\cdot\vec{r}}$ (en première approximation) ; la solution perturbée pour r grand (c.a.d. loin du diffuseur supposé localisé, ce qui permet de faire un développement limité au premier ordre de $|\vec{r}\,' - \vec{r}|$ dans l'intégrale ci-dessus) s'écrit alors (**approximation de Born de la diffusion**) :

$$f(\vec{r}) = e^{i\vec{k}_0\cdot\vec{r}} + A(\vec{k},\vec{k}_0)\, \frac{e^{ik_0 r}}{r} \quad \text{avec} \quad A(\vec{k},\vec{k}_0) = \tfrac{1}{4\pi} \iiint e^{-i(\vec{k}-\vec{k}_0)\cdot\vec{r}\,'}\, \delta k^2(\vec{r}\,')\, \mathrm{d}^3 r' \qquad (\vec{k} = k_0\hat{r}) \ ;$$

l'amplitude de diffusion est proportionnelle à la T.F. de la perturbation.

▷ Problèmes à symétrie sphérique ; séparation des variables

En coordonnées sphériques le laplacien s'écrit (*cf.* section 7.2.2)

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\,\frac{\partial}{\partial r} + \frac{1}{r^2}\,\Delta_{\theta,\varphi} \quad \text{avec} \quad \Delta_{\theta,\varphi} = \frac{1}{\sin^2\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{\partial^2}{\partial\varphi^2}\right) \ ,$$

et les harmoniques sphériques $Y_l^m(\theta,\varphi)$, introduites à la section 4.1.2, forment une base de fonctions propres de $\Delta_{\theta,\varphi}$ satisfaisant (*cf.* justification ci-dessous) :

$$\Delta_{\theta,\varphi}\, Y_l^m(\theta,\varphi) = -l\,(l+1)\, Y_l^m(\theta,\varphi) \ .$$

On cherche alors des solutions particulières de la forme $f(r,\theta,\varphi) = f_l(r)\, Y_l^m(\theta,\varphi)$. Par exemple pour $\left(\Delta + k^2(r)\right) f = 0$ on obtient :

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2}{r}\,\frac{\mathrm{d}}{\mathrm{d}r} + k^2(r) - \frac{l(l+1)}{r^2}\right) f_l(r) = 0 \ .$$

Si $k^2 = 0$ la solution est $f_l(r) = Ar^l + Br^{-(l+1)}$. Si $k^2(r) \neq 0$ n'est pas plus singulier que r^{-1} (cas de l'E.D.P. de Schrödinger pour l'atome d'hydrogène), cette expression $Ar^l + Br^{-(l+1)}$ donne le comportement de $f_l(r)$ pour $r \simeq 0$. Lorsque 0 appartient au domaine d'étude de f_l, des conditions de normalisation (ou de finitude de l'énergie) entraînent $B = 0$, ce qui équivaut à une C.L. en $r = 0$. L'ajout d'une C.L. en $r = r_0$ (cavité résonnante sphérique) ou à l'infini (état liés en mécanique quantique), conduit à la quantification des paramètres ω (pulsation) ou E (énergie) présents dans $k^2(r)$ (*cf.* section 6.4.2).

JUSTIFICATION de $\Delta_{\theta,\varphi}\, Y_l^m(\theta,\varphi) = -l\,(l+1)\, Y_l^m(\theta,\varphi)$. La définition des polynômes de Legendre par $|\vec{r}-\vec{r}_0|^{-1} = \sum_{l=0}^{\infty} \left(r_0^l\, r^{-l-1} \text{ ou } r^l\, r_0^{-l-1}\right) P_l(\cos\theta)$, leur indépendance et la propriété $\Delta(|\vec{r}-\vec{r}_0|^{-1}) = 0$ pour $\vec{r} \neq \vec{r}_0$, entraînent que $r^l\, P_l(\cos\theta)$ et $r^{-l-1}\, P_l(\cos\theta)$ sont (pour $r \neq 0$) solutions de $\Delta f = 0$. Comme $\left(\partial_r^2 + 2r^{-1}\,\partial_r\right)\, r^l = l\,(l+1)\, r^{l-2}$, on en déduit que $\Delta_{\theta,\varphi}\, P_l = -l\,(l+1)\, P_l$. Cette relation est aussi vraie pour les Y_l^m car elles se déduisent des P_l par rotation, et l'opérateur Δ est invariant par rotation.

8.3.3 Equation de Laplace $\Delta f = 0$ dans le plan et fonctions d'une variable complexe ; applications hydrodynamiques

Les fonctions $f(z)$ dérivables introduites au chapitre 2 (et associées aux transformations conformes ; *cf.* section 2.2.2) possèdent de nombreuses propriétés mathématiques remarquables (voir ouvrages spécialisés). On se contente ici de montrer qu'elles offrent des exemples simples de solutions (appelées **fonctions harmoniques**) de l'E.D.P. de Laplace $(\partial_x^2+\partial_y^2)\,V(x,y)=0$, avec des conditions aux bords correspondant à des géométries variées.

▷ Fonctions de z $f(z)=P(x,y)+iQ(x,y)$ et équations $\Delta P=\Delta Q=0$

Les parties réelle et imaginaire de $f(z)$ sont des fonctions harmoniques. Ceci est une conséquence immédiate des formules de Cauchy Riemann

$$\frac{\partial Q}{\partial x}=-\frac{\partial P}{\partial y} \quad , \quad \frac{\partial Q}{\partial y}=\frac{\partial P}{\partial x}$$

qui résultent elles mêmes de l'existence de la dérivée $f'(z)=\lim\limits_{\mathrm{d}z\to 0}\dfrac{f(z+\mathrm{d}z)-f(z)}{\mathrm{d}z}$. En effet, la limite ne devant pas dépendre de la manière dont $\mathrm{d}z$ tend vers zéro, en prenant $\mathrm{d}z=\mathrm{d}x$ et $\mathrm{d}z=i\mathrm{d}y$ on obtient $f'(z)=\partial_x P+i\partial_x Q=-i(\partial_y P+i\partial_y Q)$. On comprend heuristiquement le rôle des fonctions $f(z)$ dans la résolution de $\Delta V=0$ en remarquant que le changement de variables $(x,y)\to(z=x+iy,\overline{z}=x-iy)$ transforme cette équation en $\partial_z\partial_{\overline{z}}V=0$ et donc, V étant réel, $V=\dfrac{1}{2}\big(f(z)+\text{c.c.}\big)$.

Applications aux écoulements irrotationnels incompressibles ($\vec{v}=\overrightarrow{\nabla}P$, $\operatorname{div}\vec{v}=\Delta P=0$).
1) Pour $f(z)=z^\alpha$ $(=r^\alpha e^{i\alpha\theta})$, $\ln z$ $(=\ln r+i\theta)$ et $-i\ln z$, on a respectivement $P=r^\alpha\cos\alpha\theta$, $\ln r$ et θ. On voit que les fonctions z^α décrivent des écoulements dans des régions "en coin" $\theta\in\left[0,\dfrac{\pi}{\alpha}\right]$; en effet la vitesse est tangente aux bords car $v_\theta=r^{-1}\partial_\theta P=0$ sur les bords (radiaux) (figure 30). Les fonctions $\ln z$ et $-i\,\ln z$ décrivent des écoulements radiaux (avec source) et circulaires (*cf.* figures 26b et 26c).

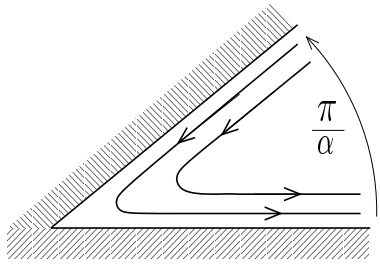


FIGURE 30

2) La fonction

$$f(z)=v_0\Big(z+\frac{r_0^2}{z}\Big)+\frac{\Gamma}{2i\pi}\,\ln\,z=P+iQ$$

décrit l'écoulement autour d'un cylindre de rayon r_0 et d'axe Oz. En effet, en explicitant $P(r,\theta)$ on obtient

$$v_r = \partial_r P = v_0\,\cos\theta\,\left(1-\frac{r_0^2}{r^2}\right)\ \text{ et }\ v_\theta = r^{-1}\partial_\theta P = -v_0\,\sin\theta\,\left(1+\frac{r_0^2}{r^2}\right)+\frac{\Gamma}{2\pi r}\ ,$$

et on voit que $\vec{v}$ est bien tangente au cylindre. On vérifie aussi que, si $|\Gamma| < 4\pi v_0 r_0$ on a $\vec{v}=0$ pour $r=r_0$ et $\sin\theta = \dfrac{\Gamma}{4\pi v_0 r_0}$ (figure 31a) et que, si $|\Gamma| > 4\pi v_0 r_0$ on a $\vec{v}=0$ pour $\theta=\dfrac{\pi}{2}$ et $r=\dfrac{\Gamma}{4\pi v_0}\pm\sqrt{\dfrac{\Gamma^2}{16\pi^2 v_0^2}-r_0^2}$ (signe + pour $\Gamma>0$ et − pour $\Gamma<0$; figure 31b).

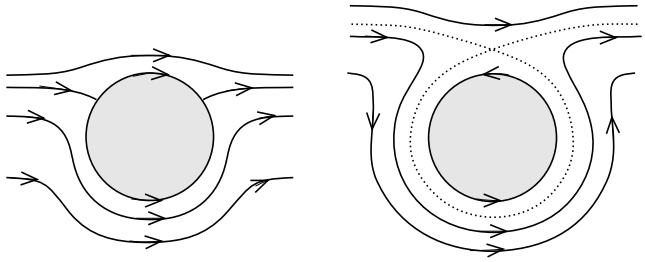


FIGURE 31

Le calcul de la force linéique $\overrightarrow{F}$ due à la pression p exercée par le fluide sur le cylindre

$$\overrightarrow{F} = \int_0^{2\pi} -p\,(\cos\theta\,\hat{x}+\sin\theta\,\hat{y})\,r_0\,\mathrm{d}\theta = -\rho v_0\Gamma\hat{y}$$

se fait en remplaçant $-p$ par $\dfrac{1}{2}\,\rho v_\theta^2$ puisque la quantité $p+\dfrac{1}{2}\,\rho v^2$ est constante (théorème de Bernouilli ; section 7.1.3) et que $v_r(r_0)=0$. On a $\overrightarrow{F}\neq 0$ si $\Gamma\neq 0$ (**effet Magnus**).

▷ Géométrie des lignes $P=$ constante et $Q=$ constante

Les formules de Cauchy Riemann montrent que, en tout point, le vecteur $\overrightarrow{\mathrm{grad}}\,Q$ se déduit de $\overrightarrow{\mathrm{grad}}\,P$ par une rotation de $\dfrac{\pi}{2}$. Les deux réseaux de courbes $P(x,y)=$ constante et $Q(x,y)=$ constante sont donc orthogonaux. Comme de plus ces gradients ont même module (égal à $|f'(z)|$), le flux de $\overrightarrow{\mathrm{grad}}\,P$ à travers toute courbe γ est égal à la circulation de $\overrightarrow{\mathrm{grad}}\,Q$ le long de γ (et inversement).

APLICATION. Pour un domaine du plan limité par les lignes $P(x,y)=P_1$ et P_2 (équipotentielles γ_1 et γ_2) et les lignes $Q(x,y)=Q_1$ et Q_2 (lignes de champ ou de courant), la quantité $\mathcal{R}$ introduite à la section 8.3.1 est donnée (pour une longueur unité perpendiculaire au plan) par la formule très simple : $\mathcal{R}=\frac{|P_2-P_1|}{|Q_2-Q_1|}$.

EXEMPLE (figure 32) : pour $f(z) = \ln\left(\frac{z}{c} + \sqrt{\frac{z^2}{c^2} - 1}\right) = P + iQ$, ce qui correspond à $\frac{z}{c} = \frac{1}{2}(e^f + e^{-f}) = \cos Q \cosh P - i \sin Q \sinh P$, les équipotentielles ($P$ fixé, Q variable) et les lignes de champ (Q fixé, P variable) forment un réseau d'ellipses et d'hyperboles homofocales dont les foyers sont situés en $z = \pm c$. Pour les équipotentielles 0 (segment $[-c, c]$) et P (ellipse) on a $\mathcal{R} = (2\pi)^{-1} \ln \frac{a+b}{c}$ (faire un tour du segment correspond à une variation de 2π de Q). On remarque aussi que $f'(z) = (z^2 - c^2)^{-\frac{1}{2}}$ est infini en $z = \pm c$; si P est un potentiel électrostatique, $|f'(z)| = |\overrightarrow{E}|$ et cette divergence du champ électrique est associée au "pouvoir des pointes".

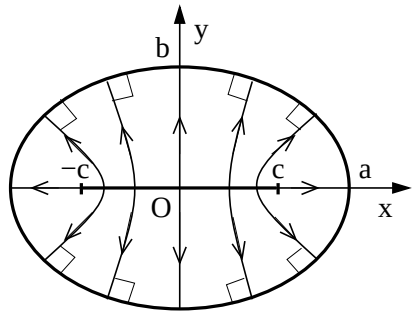


FIGURE 32

▷ Propriétés intégrales des fonctions $f(z)$

Les formules de Cauchy Riemann entraînent aussi que les champs de composantes $(P, -Q)$ et (Q, P) ont un rotationnel nul. Par application de la formule de Stokes aux parties réelle et imaginaire de l'intégrale, on obtient :

$$\oint_\gamma f(z)\,\mathrm{d}z \overset{\text{def}}{=} \oint_\gamma \left(P(x,y) + iQ(x,y)\right)\,(\mathrm{d}x + i\mathrm{d}y) = 0\ .$$

Comme on l'a vu à la section 7.2.2. cette formule n'est vraie que si γ entoure un domaine à l'intérieur duquel les rotationnels, donc $f'(z)$, sont définis. Si ce n'est pas le cas on en déduit cependant que l'intégrale (non nulle) reste inchangée lorsque γ est déformé sans qu'aucune singularité ne le traverse ; cette remarque permet souvent son calcul. Par exemple si γ fait un tour (direct) autour de z_0, en prenant le cas particulier d'un cercle $z = z_0 + re^{i\varphi}$, on obtient :

$$\oint_\gamma \frac{\mathrm{d}z}{(z - z_0)^n} = \int_0^{2\pi} i\,\frac{re^{i\varphi}}{r^n e^{in\varphi}}\,\mathrm{d}\varphi = 0 \text{ si } n \neq 1 \text{ et } 2i\pi \text{ si } n = 1$$

(cas particulier du théorème des résidus ; *cf.* ouvrages spécialisés).

Application à l'**écoulement autour d'un obstacle** (figure 33). La vitesse est donnée par $\vec{v} = \overrightarrow{\nabla} P$ ou (*cf.* Cauchy Riemann) $f'(z) = v_x - iv_y$. Si γ est la frontière orientée de l'obstacle on a :

$$\oint_\gamma f'(z)\,\mathrm{d}z = \oint_\gamma \vec{v}\,.\,\overrightarrow{\mathrm{d}l} \quad ; \quad \oint_\gamma f'^2(z)\,\mathrm{d}z = \oint_\gamma \vec{v}^2\,(\mathrm{d}x - i\mathrm{d}y)\ .$$

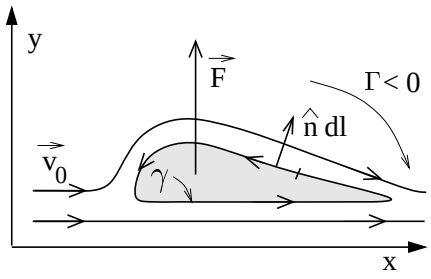


FIGURE 33

(*DÉMONSTRATION* : $\vec{v}$ étant tangent à γ, on a $v_x = v\frac{\mathrm{d}x}{\mathrm{d}l}$ et $v_y = v\frac{\mathrm{d}y}{\mathrm{d}l}$, et donc $f'(z)\,\mathrm{d}z = (v_x - iv_y)(\mathrm{d}x + i\mathrm{d}y) = v\,\mathrm{d}l$ où v est la mesure algébrique de $\vec{v}$ sur la tangente à γ orientée.)

La première intégrale donne la **circulation** Γ **du champ de vitesse** le long de γ. La seconde permet de calculer la **force linéique** $\overrightarrow{F}$ exercée par le fluide sur l'obstacle : $\overrightarrow{F} = \oint_\gamma -p\hat{n}\,\mathrm{d}l$ avec $\hat{n}\,\mathrm{d}l = (\mathrm{d}y, -\mathrm{d}x)$. En remplaçant $-p$ par $\frac{1}{2}\,\rho\,v^2$ on obtient :

$$F_x - iF_y = \oint_\gamma \tfrac{1}{2}\,\rho v^2\,(\mathrm{d}y + i\mathrm{d}x) = i\,\tfrac{\rho}{2}\,\oint_\gamma f'^2(z)\,\mathrm{d}z\ .$$

L'avantage de ces intégrales complexes est de rester inchangées si on remplace γ par tout autre circuit entourant l'obstacle. En particulier si la vitesse asymptotique est $\vec{v}_0$ (prise parallèle à $0x$), à grande distance de l'obstacle on doit avoir a priori

$$f'(z) = v_0 + a_{-1}\,z^{-1} + a_{-2}\,z^{-2}\ldots \quad \text{et} \quad a_{-1} = \tfrac{\Gamma}{2i\pi}$$

(*cf.* intégrale de $f'(z)$). En considérant le terme en z^{-1} dans $f'^2(z)$ on voit que $F_x - iF_y = i\rho v_0\Gamma$, soit $F_x = 0$ et $F_y = -\rho v_0\Gamma$ ou vectoriellement $\overrightarrow{F} = \rho\vec{v}_0 \wedge \overrightarrow{\Gamma}$. Ce résultat général montre la nécessité de la circulation de l'air autour d'une aile pour avoir un effet de **portance**.

REMARQUES. 1) L'intérêt principal des variables complexes est, partant d'un écoulement connu, de déduire d'autres écoulements en déformant l'obstacle à l'aide d'une **transformation conforme** arbitraire $Z = g(z)$. Dans le plan (X, Y) la fonction $F(Z) = f(z)$ correspond alors à un écoulement dont les équipotentielles et les lignes de champ sont les transformées de leurs homologues associées à $f(z)$ dans le plan (x, y) (*cf.* ouvrages spécialisés).

2) Un écoulement stationnaire incompressible avec viscosité vérifie $-\overrightarrow{\mathrm{grad}}p + \rho\vec{g} + \eta\,\Delta\vec{v} = 0$ (*cf.* section 7.2.4). En posant $v_x = \partial_y\Psi$, $v_y = -\partial_x\Psi$ et en prenant le rotationnel des deux membres de l'équation on obtient $\Delta^2\Psi = 0$ ou $\partial_z^2\,\partial_{\overline{z}}^2\,\Psi = 0$. Ceci conduit à des solutions du type $\Psi = f(z) + \overline{z}\,g(z) + \text{c.c.}$. Ces **fonctions biharmoniques** se rencontrent aussi en élasticité, dans des problèmes plans de déformations.

Chapitre 9

Principes variationnels

Les principes d'extremum ont été introduits de longue date, en optique avec le principe de Fermat, et en mécanique avec les **principes de moindre action** de Maupertuis, Lagrange et Hamilton. Ils permettent de caractériser de manière élégante les trajectoires ou évolutions réellement observées par rapport à d'autres a priori possibles et satisfaisant les mêmes conditions aux limites. Ils conduisent également à une technique de calcul, les **équations d'Euler-Lagrange**, très efficace et permettant n'importe quel choix de coordonnées. Enfin ils constituent un moyen remarquable d'associer les grandes **lois de conservation** (L.C.) de la physique à des propriétés d'**invariance par symétrie**.

9.1 EXEMPLES HISTORIQUES ; FORMALISMES DE LAGRANGE ET DE HAMILTON

Le tableau ci-dessous présente quatre principes d'extremum que nous commençons par commenter et relier entre eux en admettant provisoirement les équations d'Euler-Lagrange établies à la section 9.1.3 et rappelées dans le tableau.

9.1.1 Principes de Fermat, Maupertuis, Lagrange

En paramétrant le chemin de A à B par $x(\lambda)$, $y(\lambda)$ et $z(\lambda)$ le **principe de Fermat** s'écrit "$\int_0^1 n(x,y,z)\sqrt{x'^2+y'^2+z'^2}\,\mathrm{d}\lambda$ extremum" comme celui de Lagrange. Le signe $'$ désigne la dérivation par rapport à λ qui joue le rôle du temps et le lagrangien est :

$$L(x,y,z,x',y',z') = n(x,y,z)\sqrt{x'^2+y'^2+z'^2}\ .$$

L'E.D. d'Euler-Lagrange pour la variable $q_i = x$ (*cf.* tableau)

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial x'} = \frac{\partial L}{\partial x} \quad \text{avec} \quad \frac{\partial L}{\partial x} = \frac{\partial n}{\partial x}\sqrt{x'^2+y'^2+z'^2} \quad \text{et} \quad \frac{\partial L}{\partial x'} = n\frac{x'}{\sqrt{x'^2+y'^2+z'^2}} = nu_x$$

(où $\hat{u} = \dfrac{\mathrm{d}\vec{r}}{\mathrm{d}l}$ est le vecteur unitaire tangent au rayon) conduit immédiatement à la **loi de Descartes** $\dfrac{\mathrm{d}}{\mathrm{d}l}(n\hat{u}) = \overrightarrow{\mathrm{grad}}\, n$ projetée sur l'axe x.

Mathématiquement identique au principe de Fermat, le **principe de Maupertuis** s'applique à des particules d'énergie totale E fixée et d'énergie potentielle $U(\vec{r})$, donc en particulier à l'**optique électronique**. Le module $p(\vec{r})$ de la quantité de mouvement en tout point est déterminé à partir de la L.C. de l'énergie $E = E_{cin} + U(\vec{r})$ (par exemple $p(\vec{r}) = \sqrt{2m(E - U(\vec{r}))}$ si $E_{cin} = \dfrac{\vec{p}^{\,2}}{2m}$). Comme $\mathrm{d}E_{cin} = v\,\mathrm{d}p = -\mathrm{d}U$, relation également vraie en relativité d'Einstein, l'équation $\dfrac{\mathrm{d}\vec{p}}{\mathrm{d}l} = \overrightarrow{\mathrm{grad}}\, p$ multipliée par $v = \dfrac{\mathrm{d}l}{\mathrm{d}t}$ donne la **loi de Newton** $\dfrac{\mathrm{d}\vec{p}}{\mathrm{d}t} = -\overrightarrow{\mathrm{grad}}\, U(\vec{r})$.

REMARQUE. **Lien avec les ondes**. Comme en optique ondulatoire le module du vecteur d'onde vérifie $k(\vec{r}) = \frac{\omega}{c} n(\vec{r})$ (*cf.* section 7.3.2), le principe de Fermat s'écrit aussi $\int_A^B k(\vec{r})\,\mathrm{d}l$ extremum. Il s'étend, sous cette forme, à d'autres ondes (ultrasons, ondes sismiques, *etc.*). La **relation de de Broglie** $\vec{p} = \hbar\vec{k}$ unifie les principes de Fermat et de Maupertuis.

Principes	Quantités extremum	Equations	Lois
Fermat	$\int_A^B n(\vec{r})\,\mathrm{d}l$	$\dfrac{\mathrm{d}}{\mathrm{d}l}(n\hat{u}) = \overrightarrow{\mathrm{grad}}\, n$	Descartes
Maupertuis	$\int_A^B p(\vec{r})\,\mathrm{d}l$	$\dfrac{\mathrm{d}}{\mathrm{d}l}\vec{p} = \overrightarrow{\mathrm{grad}}\, p$	Newton (énergie fixée)
Lagrange ex : $\underline{q} = \vec{r}\ ;\ \underline{\dot{q}} = \vec{v}$	$S = \int_{\underline{q}_1 t_1}^{\underline{q}_2 t_2} L(\underline{q}, \underline{\dot{q}}, t)\,\mathrm{d}t$ $L = \dfrac{1}{2}m\vec{v}^{\,2} - U(\vec{r}, t)$	$\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial L}{\partial \dot{q}_i}\right) = \dfrac{\partial L}{\partial q_i}$ $\dfrac{\mathrm{d}}{\mathrm{d}t}m\vec{v} = -\overrightarrow{\mathrm{grad}}\, U$	Equations d' Euler-Lagrange Newton
Hamilton ex : $\underline{q} = \vec{r}\ ;\ \underline{p} = \vec{p}$	$\int_{\underline{q}_1 \underline{p}_1 t_1}^{\underline{q}_2 \underline{p}_2 t_2} (\underline{p}\cdot\underline{\dot{q}} - H(\underline{q}, \underline{p}, t))\,\mathrm{d}t$ $H = \dfrac{\vec{p}^{\,2}}{2m} + U(\vec{r}, t)$	$\dot{q}_i = \dfrac{\partial H}{\partial p_i}\ ;\ \dot{p}_i = -\dfrac{\partial H}{\partial q_i}$ $\vec{v} = \dfrac{\vec{p}}{m}\ ;\ \dfrac{\mathrm{d}\vec{p}}{\mathrm{d}t} = -\overrightarrow{\mathrm{grad}}\, U$	Equations de Hamilton Newton

▷ Principe de Lagrange ; propriétés générales

Ce principe est la forme la plus générale sous laquelle on peut exprimer un problème d'extremum impliquant des fonctions $q_i(t)$ et leurs dérivées premières $\dot{q}_i(t)$. $\underline{q}(t)$ est un chemin dans l'**espace de configuration** $(q_1, \cdots, q_n)$ dont les extrémités $\underline{q}_1$ et $\underline{q}_2$ aux instants t_1 et t_2 sont fixées. On discute ici ses principales propriétés.

Addition au lagrangien d'une dérivée totale. Comme les extrémités sont fixées, les équations d'Euler-Lagrange sont inchangées si on ajoute au lagrangien $L(\underline{q}, \underline{\dot{q}}, t)$ une dérivée totale par rapport au temps $\dfrac{\mathrm{d}}{\mathrm{d}t} F(\underline{q}(t), t)$. En effet son intégrale $F(\underline{q}(t_2), t_2) - F(\underline{q}(t_1), t_1)$ a une variation nulle.

Changements de variables. Le formalisme lagrangien est invariant vis à vis des changements de variables. En effet si on pose $\underline{q} = \underline{f}(\underline{Q}, t)$, la quantité à extrémaliser devient $\int L\big(\underline{f}(\underline{Q},t), \frac{\mathrm{d}}{\mathrm{d}t}\underline{f}(\underline{Q},t), t\big)\,\mathrm{d}t = \int \mathcal{L}(\underline{Q}, \dot{\underline{Q}}, t)\,\mathrm{d}t$ et conduit aux équations :

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\underline{Q}}} = \frac{\partial \mathcal{L}}{\partial \underline{Q}} \quad \text{équivalentes à} \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\underline{q}}} = \frac{\partial L}{\partial \underline{q}} \ .$$

Symétries. Si un problème de mécanique est invariant par rotation autour d'un axe Oz, il est judicieux d'écrire le lagrangien en coordonnées cylindriques par exemple :

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) - U(r,\theta,z) \ .$$

Comme U ne doit pas dépendre de θ, l'équation $\dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial \mathcal{L}}{\partial \dot{\theta}} = \dfrac{\partial \mathcal{L}}{\partial \theta} = 0$ conduit à la conservation de la composante z du moment cinétique $mr^2\dot{\theta} = \dfrac{\partial \mathcal{L}}{\partial \dot{\theta}}$. Si en plus U ne dépend pas de z, on a conservation de la composante z de la quantité de mouvement $m\dot{z} = \dfrac{\partial \mathcal{L}}{\partial \dot{z}}$. De façon générale si Q_i n'apparaît pas dans le lagrangien $\mathcal{L}$ (on dit que Q_i est une **coordonnée cyclique**), la quantité $\dfrac{\partial \mathcal{L}}{\partial \dot{Q}_i}$ est une constante (**intégrale première**) du mouvement.

Contraintes. Si il existe des relations entre les coordonnées q_i, n'impliquant pas leurs dérivées $\dot{q}_i$, on exprime le lagrangien en fonction de variables indépendantes et de leurs dérivées. Par exemple pour un pendule dont le point de fixation oscille verticalement (figure 1),

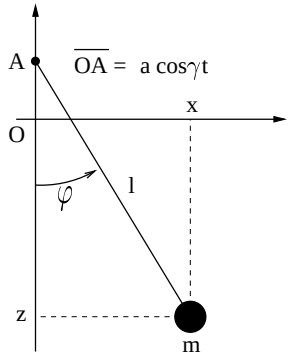


FIGURE 1

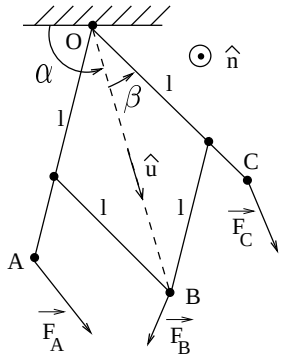


FIGURE 2

$z = a\cos\gamma t - l\cos\varphi$ et $x = l\sin\varphi$ et le lagrangien $L = \dfrac{1}{2}m(\dot{x}^2 + \dot{z}^2) - mgz$ s'écrit $\mathcal{L}(\varphi, \dot{\varphi}, t) = \dfrac{1}{2}ml^2\dot{\varphi}^2 + ml(g - a\gamma^2\cos\gamma t)\cos\varphi$ (après élimination de dérivées totales). L'équation d'Euler-Lagrange linéarisée $\ddot{\varphi} + \dfrac{g}{l}\Big(1 - \dfrac{a\gamma^2}{g}\cos\gamma t\Big)\varphi = 0$ décrit un oscillateur paramétrique (*cf.* sections 6.1.2 et 6.4.4).

Cas des **équilibres**. Pour un lagrangien du type "(énergie cinétique) $-(U(\vec{r}_1,\vec{r}_2,\cdots,\vec{r}_n)$ énergie potentielle)", le principe d'extremum pour une position d'équilibre se ramène à $\delta U=0$ puisque toute variation de l'énergie cinétique est du second ordre. Si les $\vec{r}_i$ sont indépendants cela donne les conditions connues $\partial_{\vec{r}_i}U=0$, mais si ils ne le sont pas on a là un moyen, appelé historiquement **théorème des travaux virtuels**, d'obtenir l'équilibre. Ainsi lorsque des forces extérieures constantes $\vec{F}_i$ sont appliquées à des points $\vec{r}_i$ d'un système articulé l'équilibre est déterminé par

$$\delta U=\delta\left(-\textstyle\sum_i \vec{F}_i\cdot\vec{r}_i\right)=-\sum_i \vec{F}_i\cdot\delta\vec{r}_i=0\ ,$$

les $\vec{r}_i$ étant exprimés en fonction des variables libres du système. Prenons l'exemple d'un losange articulé (figure 2), avec des branches sans masse et des forces $\vec{F}_A$, $\vec{F}_B$, $\vec{F}_C$ appliquées en A, B, C (prenant en compte la pesanteur si A, B et C sont massifs). En fonction des variables libres α et β, on a $\delta\vec{r}_A=(\delta\alpha-\delta\beta)\,\hat{n}\wedge\vec{OA}$, $\delta\vec{r}_B=\delta\alpha\,\hat{n}\wedge\vec{OB}+\delta(2l\,\cos\beta)\hat{u}$ et $\delta\vec{r}_C=(\delta\alpha+\delta\beta)\,\hat{n}\wedge\vec{OC}$. On vérifie qu'annuler le facteur de $\delta\alpha$ dans δU revient à exprimer que le moment résultant des forces par rapport à O est nul ; annuler le facteur de $\delta\beta$ conduit à la relation moins triviale $2l\,\sin\beta\,\vec{F}_B\cdot\hat{u}=(\vec{OC}\wedge\vec{F}_C-\vec{OA}\wedge\vec{F}_A)\cdot\hat{n}$.

9.1.2 Principe de Hamilton dans l'espace de phase ($\underline{q}$,$\underline{p}$)

Dans ce principe les variables $q_i(t)$ et $p_i(t)$ sont par définition indépendantes et

$$\sum_i p_i\dot{q}_i-H(q_1,q_2\cdots q_n;p_1,p_2\cdots p_n;t)=L_H(q_1,q_2\cdots q_n;p_1,p_2\cdots p_n;\dot{q}_1,\dot{q}_2\cdots\dot{q}_n;t)$$

est le lagrangien relatif à ces $2n$ variables ; $H(\underline{q},\underline{p},t)$ est appelé **hamiltonien**. On vérifie que les E.D. d'Euler-Lagrange $\dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial L_H}{\partial\dot{q}_i}\Big)=\dfrac{\partial L_H}{\partial q_i}$ et $\dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial L_H}{\partial\dot{p}_i}\Big)=\dfrac{\partial L_H}{\partial p_i}$ sont les **E.D. de Hamilton** :

$$\dot{p}_i=-\frac{\partial H}{\partial q_i}\quad\text{et}\quad 0=\dot{q}_i-\frac{\partial H}{\partial p_i}\ .$$

Plus généralement l'évolution de toute fonction $g(\underline{q},\underline{p})$ s'écrit aussi $\dot{g}=-\{H,g\}$ en introduisant les **crochets de Poisson** $\{f,g\}=\sum_i\big(\partial_{q_i}f\,\partial_{p_i}g-\partial_{q_i}g\,\partial_{p_i}f\big)$. Une propriété remarquable du mouvement dans l'espace de phase est de conserver le volume (**théorème de Liouville**). Ceci résulte de ce que la "divergence" (à $2n$ dimensions) du champ de vitesse $(\dot{\underline{q}},\dot{\underline{p}})$ est nulle : $\sum_i\big(\partial_{q_i}(\partial_{p_i}H)+\partial_{p_i}(-\partial_{q_i}H)\big)=0$.

REMARQUE. **Lien avec les ondes**. En identifiant $\vec{p}$ à $\hbar\vec{k}$ et H à $\hbar\omega$ le principe de Hamilton devient :

$$\Phi=\int\Big[\vec{k}(\vec{r},t)\cdot\mathrm{d}\vec{r}-\omega(\vec{k},\vec{r},t)\,\mathrm{d}t\Big]\quad\text{extremum}\ .$$

L'interprétation physique est la suivante : Φ désigne une variation de phase et la fonction $\omega(\vec{k},\vec{r},t)$ est la relation de dispersion satisfaite par l'onde en $\vec{r}$ à l'instant t. Les équations de Hamilton

$$\tfrac{\mathrm{d}\vec{r}}{\mathrm{d}t}=\tfrac{\partial\omega}{\partial\vec{k}}\ \ (=\ \textbf{vitesse de groupe}\ \text{locale})\quad\text{et}\quad\tfrac{\mathrm{d}\vec{k}}{\mathrm{d}t}=-\tfrac{\partial\omega}{\partial\vec{r}}$$

définissent le mouvement $\vec{r}(t)$ et les caractéristiques $\vec{k}(t)$ et $\omega(t)$ d'un **paquet d'ondes**. L'invariance du volume $\mathrm{d}^3\vec{r}\,\mathrm{d}^3\vec{k}$ (théorème de Liouville) équivaut à celle de l'**étendue optique** vue à la section 3.2.5.

▷ Comparaison des approches lagrangienne et hamiltonienne

De manière générale on passe de la première à la seconde en posant

$$p_i=\frac{\partial L}{\partial\dot{q}_i}\quad(p_i\ \textbf{moment conjugué}\ \text{de}\ q_i)\quad\text{et}\quad H(\underline{q},\underline{p},t)=\sum_i p_i\dot{q}_i-L(\underline{q},\dot{\underline{q}},t)\ ,$$

étant entendu que pour écrire H il faut exprimer les $\dot{q}_i$ en fonction des q_i, p_i et t, par exemple : $L = \frac{1}{2}m\vec{v}^2 - U \Longrightarrow \vec{p} = m\vec{v} \Longrightarrow H = \vec{p}\cdot\vec{v} - L = \frac{\vec{p}^2}{2m} + U$. Les E.D. de Hamilton découlent alors de $\mathrm{d}H = \sum_i \big(p_i \mathrm{d}\dot{q}_i + \dot{q}_i \mathrm{d}p_i - \frac{\partial L}{\partial q_i}\,\mathrm{d}q_i - \frac{\partial L}{\partial \dot{q}_i}\,\mathrm{d}\dot{q}_i\big) - \frac{\partial L}{\partial t}\,\mathrm{d}t$. En effet, puisque $p_i = \frac{\partial L}{\partial \dot{q}_i}$, on a bien $\frac{\partial H}{\partial p_i} = \dot{q}_i$ et $\frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i} = -\frac{\mathrm{d}}{\mathrm{d}t}\big(\frac{\partial L}{\partial \dot{q}_i}\big) = -\dot{p}_i$. On remarque aussi que au cours du mouvement, c'est à dire lorsque $\dot{p}_i = \frac{\partial L}{\partial q_i}$, on a $\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{\partial L}{\partial t}$; donc H est une constante du mouvement lorsque L ne dépend pas explicitement du temps.

En mécanique classique l'intérêt du formalisme de Hamilton est d'être invariant dans des changements de coordonnées, dits canoniques, pouvant mélanger les $2n$ variables p_i et q_i et préservant les crochets de Poisson (voir section 9.1.4), ce qui le rend rend encore plus souple que le formalisme de Lagrange. L'autre intérêt est de permettre un passage du hamiltonien classique au **hamiltonien quantique** (presque automatique) en remplaçant les p_i par les opérateurs $-i\hbar\partial_{q_i}$, et les crochets de Poisson par des commutateurs (*cf.* section 4.4.1).

9.1.3 Equations d'Euler-Lagrange ; symétries et lois de conservation ; E.D.P. d'Hamilton-Jacobi

▷ Variation générale de l'action et E.D. d'Euler-Lagrange

On peut écrire la quantité à extrémaliser sous la forme

$$S = \int_{-\infty}^{\infty} L(\underline{q}, \underline{\dot{q}}, t)\,\chi(t)\,\mathrm{d}t$$

en introduisant la fonction caractéristique $\chi(t)$ de l'intervalle $[t_1, t_2]$ (figure 3). Dans une variation infinitésimale qui concerne à la fois les fonctions $q_i(t)$ et l'intervalle $[t_1, t_2]$ donc $\chi(t)$,

$$q_i(t) \to q_i(t) + \delta q_i(t) \Longrightarrow \dot{q}_i(t) \to \dot{q}_i(t) + \frac{\mathrm{d}}{\mathrm{d}t}\delta q_i(t) \quad \text{et} \quad \chi(t) \to \chi(t+\epsilon(t)) = \chi(t) + \epsilon(t)\,\frac{\mathrm{d}\chi}{\mathrm{d}t}\,,$$

la variation de l'action contient deux contributions, une liée à la variation du lagrangien $\delta L = \sum_i \Big(\frac{\partial L}{\partial q_i}\,\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\,\delta\dot{q}_i\Big)$ et une à la variation de χ. Après une intégration par parties on obtient :

$$\delta S = \int_{-\infty}^{\infty} \chi(t) \sum_i \Big[\frac{\partial L}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big)\Big]\,\delta q_i(t)\,\mathrm{d}t + \int_{-\infty}^{\infty} \frac{\mathrm{d}\chi}{\mathrm{d}t}\Big[\epsilon L - \sum_i \frac{\partial L}{\partial \dot{q}_i}\,\delta q_i\Big]\,\mathrm{d}t = \delta S_1 + \delta S_2\,.$$

Supposons t_1 et t_2 fixés ($\epsilon = 0$) ainsi que les extrémités des chemins ($\delta q_i(t_1) = \delta q_i(t_2) = 0$). Cela entraîne $\delta S_2 = 0$ puisque le support de $\frac{\mathrm{d}\chi}{\mathrm{d}t}$ se limite aux instants t_1 et t_2 (figure 3). Alors le principe d'extremum exprime que pour le mouvement réellement observé $\delta S_1 = 0$ quelles que soient les variations $\delta q_i(t)$. On en déduit les **E.D. d'Euler-Lagrange** :

$$\frac{\partial L}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big) = \dot{p}_i \ (p_i \text{ moment conjugué de } q_i).$$

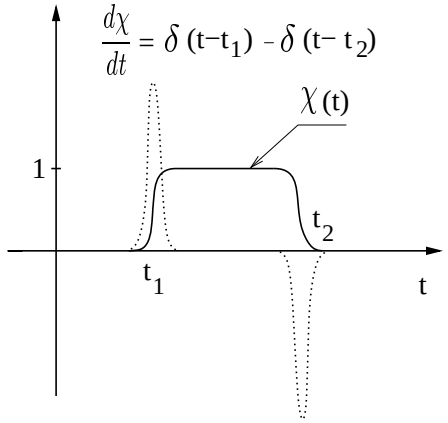


FIGURE 3

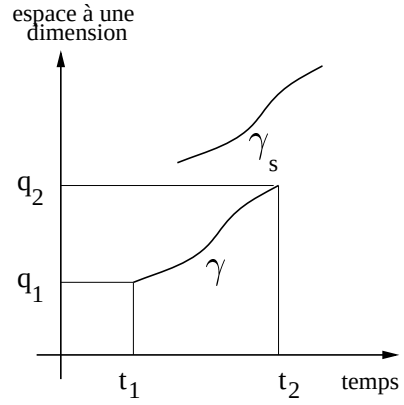


FIGURE 4

▷ Symétries et lois de conservation

On appelle symétrie une opération qui fait passer d'un chemin γ quelconque (a priori non extrémal) à un chemin γ_s sans modifier l'action (figure 4) :

$$S(\gamma) = S(\gamma_s)\ .$$

Dans la suite on ne considère que des opérations infinitésimales.

EXEMPLES : l'opération $\chi(t) \to \chi(t+\epsilon)$ et $q_i(t) \to q_i(t+\epsilon) = q_i(t) + \epsilon\,\dot{q}_i(t)$ avec ϵ constant est une symétrie si L ne dépend pas explicitement du temps (**stationnarité** de L) ; en effet : $\displaystyle\int_{-\infty}^{\infty} L\big(\underline{q}(t+\epsilon), \underline{\dot{q}}(t+\epsilon)\big)\,\chi(t+\epsilon)\,\mathrm{d}t = \int_{-\infty}^{\infty} L\big(\underline{q}(t), \underline{\dot{q}}(t)\big)\,\chi(t)\,\mathrm{d}t$. Autres exemples avec χ inchangée ($\epsilon = 0$) : si $L(\vec{r}_i, \vec{v}_i)$, lagrangien de n particules, ne dépend de leurs positions que par l'intermédiaire des différences $\vec{r}_i - \vec{r}_j$, alors l'opération $\vec{r}_i \to \vec{r}_i + \vec{\epsilon}$ est une **symétrie de translation** car $L(\vec{r}_i + \vec{\epsilon}, \vec{v}_i) = L(\vec{r}_i, \vec{v}_i)$; de même si $L(\vec{r}_i, \vec{v}_i)$ ne fait intervenir que des produits scalaires des $\vec{r}_i$ et $\vec{v}_i$, l'opération $\vec{r}_i \to \vec{r}_i + \delta\theta\,\vec{n} \wedge \vec{r}_i$, $\vec{v}_i \to \vec{v}_i + \delta\theta\,\vec{n} \wedge \vec{v}_i$ est une **symétrie de rotation**.

Considérons l'action écrite pour le chemin γ correspondant au mouvement réel et effectuons une opération de symétrie infinitésimale. Dans cette opération $\delta S = 0$ par définition d'une symétrie, et $\delta S_1 = 0$ puisque les équations d'Euler-Lagrange sont satisfaites sur γ. Par conséquent

$$\delta S_2 = \int_{-\infty}^{\infty} \frac{\mathrm{d}\chi}{\mathrm{d}t}\left[\epsilon L - \sum_i \frac{\partial L}{\partial \dot{q}_i}\,\delta q_i\right]\mathrm{d}t \ = 0 = \int_{-\infty}^{\infty} \chi\,\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_i \frac{\partial L}{\partial \dot{q}_i}\,\delta q_i - \epsilon L\right]\mathrm{d}t$$

(intégration par parties). Comme le support de χ est arbitraire on en déduit que sur la trajectoire réelle la quantité $\displaystyle\sum_i p_i\,\delta q_i - \epsilon L$ est conservée. Pour les trois exemples ci-dessus on a respectivement $\delta q_i = \epsilon \dot{q}_i$, $\delta \vec{r}_i = \vec{\epsilon}$ et $\epsilon = 0$, $\delta \vec{r}_i = \delta\theta\,\vec{n} \wedge \vec{r}_i$ et $\epsilon = 0$, ce qui conduit à :

- stationnarité de $L \Longrightarrow$ **L.C. de l'énergie** $H = \sum_i p_i \dot{q}_i - L$,

- homogénéité de $L \Longrightarrow$ **L.C. de la quantité de mouvement** $\vec{p} = \sum_i \dfrac{\partial L}{\partial \vec{v}_i}$,

- isotropie de $L \Longrightarrow$ **L.C. du moment cinétique** $\vec{J} = \sum_i \vec{r}_i \wedge \dfrac{\partial L}{\partial \vec{v}_i}$.

▷ Différentielle de l'action ; E.D.P. d'Hamilton-Jacobi

Soit $S(\underline{q}_2, t_2; \underline{q}_1, t_1)$ la valeur de l'action S pour la trajectoire allant de $\underline{q}_1$ à $\underline{q}_2$ dans l'intervalle de temps $[t_1, t_2]$. On considère la trajectoire voisine qui va de $\underline{q}_1 + \mathrm{d}\underline{q}_1$ à $\underline{q}_2 + \mathrm{d}\underline{q}_2$ dans l'intervalle $[t_1 + \mathrm{d}t_1, t_2 + \mathrm{d}t_2]$. En reprenant l'expression ci-dessus de δS_2 avec $\epsilon(t_{1,2}) = -\mathrm{d}t_{1,2}$ et (réflexion laissée au lecteur, *cf.* figure 4) $\delta\underline{q}(t_{1,2}) = \mathrm{d}\underline{q}_{1,2} - \dot{\underline{q}}_{1,2}\,\mathrm{d}t_{1,2}$, on obtient la différentielle

$$\mathrm{d}S = (\underline{p}_2\,\mathrm{d}\underline{q}_2 - H_2\,\mathrm{d}t_2) - (\underline{p}_1\,\mathrm{d}\underline{q}_1 - H_1\,\mathrm{d}t_1)$$

avec $\underline{p}_{1,2} = \underline{p}(t_{1,2})$ et $H_{1,2} = H(\underline{q}_{1,2}, \underline{p}_{1,2}, t_{1,2})$. On en déduit que, en temps que fonction de $(\underline{q}_2, t_2) \equiv (\underline{q}, t)$, S vérifie l'**E.D.P. de Hamilton-Jacobi** :

$$\frac{\partial S}{\partial t} = -H\Big(\underline{q}, \frac{\partial S}{\partial \underline{q}}, t\Big) \ .$$

Par exemple pour une particule libre galiléenne pour laquelle $L = \frac{1}{2} m \overrightarrow{v}^2$ et $H = \frac{1}{2m}\vec{p}^2$, on a :

$$S(\overrightarrow{r}_2, t_2; \overrightarrow{r}_1, t_1) = \frac{1}{2} m \frac{(\overrightarrow{r}_2 - \overrightarrow{r}_1)^2}{t_2 - t_1} \quad ; \quad \frac{\partial S}{\partial t} = -\frac{1}{2m}\,(\overrightarrow{\nabla S})^2 \ .$$

On remarquera que $\mathrm{d}S$ ci-dessus généralise la différentielle du chemin optique $\mathrm{d}L = n_B \widehat{u}_B \cdot \mathrm{d}\overrightarrow{r}_B - n_A \widehat{u}_A \cdot \mathrm{d}\overrightarrow{r}_A$ vue à la section 3.2.3 et que l'E.D.P. de Hamilton-Jacobi généralise l'**E.D.P. eikonale** $|\overrightarrow{\nabla L}| = n(\overrightarrow{r})$ de la section 7.3.2.

9.1.4 Equations de Hamilton et géométrie de l'espace de phase

▷ Espace de phase ; géométrie symplectique

Rappelons que, à une dimension, un vecteur de l'espace de phase pour une particule est défini par les coordonnées (q, p) ($q \equiv x$ position et $p \equiv m\dot{x}$ quantité de mouvement) et par $(\overrightarrow{r}_1, \cdots, \overrightarrow{r}_N, \overrightarrow{p}_1, \cdots, \overrightarrow{p}_N)$ pour N particules à trois dimensions. Plus généralement, on note $\underline{x}$ le vecteur colonne de coordonnées $(q_1, \cdots, q_n, p_1, \cdots, p_n)$. La **géométrie symplectique** est définie par l'ensemble des transformations linéaires $\underline{x} \to \underline{x}' = \mathbf{M}\underline{x}$ qui laissent invariante la forme antisymétrique (*cf.* section 4.1.3), dite symplectique :

$$\sigma(\underline{x}, \underline{y}) = \underline{x}^t \mathbf{J} \underline{y} \quad \text{avec } \mathbf{J} = \begin{pmatrix} 0 & \mathbf{I}_n \\ -\mathbf{I}_n & 0 \end{pmatrix} \text{ et } \mathbf{I}_n \text{ matrice identité } n \times n \ .$$

Les **matrices symplectiques** $\mathbf{M}$ vérifient $\mathbf{M}^t\mathbf{J}\mathbf{M} = \mathbf{J}$. Par exemple si $n = 1$, $\underline{x} = (q^{(1)}, p^{(1)})^t$ et $\underline{y} = (q^{(2)}, p^{(2)})^t$, on a $\sigma(\underline{x}, \underline{y}) = q^{(1)}p^{(2)} - q^{(2)}p^{(1)}$ et les matrices $\mathbf{M}$ sont les matrices 2×2 de déterminant égal à 1. Plus généralement $\sigma(\underline{x}, \underline{y}) = \sum_i (q_i^{(1)} p_i^{(2)} - q_i^{(2)} p_i^{(1)})$ est la somme des aires algébriques des projections sur les "plans" (q_i, p_i) du parallélogramme construit sur $\underline{x}$ et $\underline{y}$ et les matrices $\mathbf{M}$ conservent cette somme.

REMARQUE. A l'aide de la relation $\mathbf{J}^t = -\mathbf{J}$, on établit facilement qu'une matrice symplectique infinitésimale s'écrit $\mathbf{M} = \mathbf{I} + \epsilon\mathbf{J}\mathbf{S}$ avec $\mathbf{S} = \mathbf{S}^t$, matrice symétrique.

▷ Equations de Hamilton et transformations canoniques

Le lien de la géométrie symplectique avec la physique se fait via les **équations de Hamilton**

$$\dot{q}_i = \partial_{p_i} H \quad \text{et} \quad \dot{p}_i = -\partial_{q_i} H \ ,$$

conséquences de l'existence d'un principe variationnel (*cf.* section 9.1.2). On vérifie en effet que ces équations s'écrivent

$$\dot{\underline{x}} = \mathbf{J}\partial_{\underline{x}} H \quad \text{avec} \quad \partial_{\underline{x}} \equiv (\partial_{q_1}, \cdots, \partial_{q_n}, \partial_{p_1}, \cdots, \partial_{p_n})^t \ ,$$

et qu'elles sont laissées invariantes par les transformations ci-dessus : $\dot{\underline{x}}' = \mathbf{MJ}\partial_{\underline{x}} H = \mathbf{MJM}^t \partial_{\underline{x}'} H$ (*cf.* section 7.1.4) $= \mathbf{J}\partial_{\underline{x}'} H$. Plus généralement elles sont laissées invariantes par les **transformations canoniques** $(\underline{q}, \underline{p}) \to (\underline{q}'(\underline{q}, \underline{p}), \underline{p}'(\underline{q}, \underline{p}))$ caractérisées par une matrice jacobienne $\frac{\partial(\underline{q}', \underline{p}')}{\partial(\underline{q}, \underline{p})}$ symplectique.

On vérifie aussi que l'évolution d'une perturbation infinitésimale est symplectique. En effet on a

$$\delta\dot{\underline{x}} = \mathbf{JH}'' \, \delta\underline{x} \quad \text{ou} \quad \delta\underline{x}(t + \mathrm{d}t) = \big(\mathbf{I} + \mathbf{JH}'' \, \mathrm{d}t\big) \, \delta\underline{x}(t)$$

avec $\mathbf{H}'' = [\mathbf{H}'']^t$ **matrice hessienne** (des dérivées secondes) du hamiltonien H.

REMARQUE. Les lois de l'**optique géométrique**, également issues d'un principe variationnel (de Fermat), mettent aussi en jeu des transformations symplectiques (*cf.* les matrices de transfert de déterminant égal à 1 des sections 4.3.2 et 6.4.1).

9.2 PRINCIPES DE MOINDRE ACTION ET GÉNÉRALISATION DES MOUVEMENTS INERTIELS

Un éclairage intéressant sur les principes d'extremum en mécanique est fourni par un retour sur les mouvements inertiels qui jouent un rôle essentiel en physique. On rappelle (*cf.* section 4.3.3) que le **mouvement inertiel**, c'est-à-dire à vitesse uniforme entre deux points d'espace-temps (A, t_1) et (B, t_2) maximise le **temps propre :**

$$\tau = \int_{(A,t_1)}^{(B,t_2)} \sqrt{1 - \frac{v^2}{c^2}} \, \mathrm{d}t \ .$$

9.2.1 Collisions et introduction de la masse inertielle

Si on considère le choc élastique de deux particules (figure 5a) entre des positions initiales (A et B) et finales (A' et B') fixées, on sait que la particule la plus massive ($m_1 > m_2$) est celle dont le mouvement est le plus proche d'un mouvement inertiel (ACA' plus "proche" de AA' que BCB' de BB'). Il est donc naturel d'introduire les masses par le principe "$m_1\tau_1 + m_2\tau_2$ maximum", où chaque temps propre τ est pondéré par la masse m. Il est remarquable que, plus généralement, les L.C. de la quantité de mouvement et de l'énergie sont une conséquence du principe très simple

$$\sum_i m_i \tau_i \quad \text{extremum} \ .$$

DÉMONSTRATION : considérons l'exemple "typique" $1 + 2 \to 3 + 4 + 5$, les particules finales pouvant différer des particules initiales (figure 5b). Entre les instants initial ou final et l'instant du choc, chaque particule i effectue à vitesse uniforme la translation $\vec{R}_i$ dans le temps T_i, avec $T_1 = T_2$ avant le choc et $T_3 = T_4 = T_5$ après. Pour chaque particule i on a $\tau_i = \left(T_i^2 - \frac{R_i^2}{c^2}\right)^{\frac{1}{2}}$ et

$$m_i c^2 \mathrm{d}\tau_i = m_i c^2 \Big(T_i \, \mathrm{d}T_i - \frac{\vec{R}_i \cdot \mathrm{d}\vec{R}_i}{c^2}\Big)\Big(T_i^2 - \frac{R_i^2}{c^2}\Big)^{-\frac{1}{2}} = E_i \, \mathrm{d}T_i - \vec{p}_i \cdot \mathrm{d}\vec{R}_i \ ,$$

où $E_i = m_i c^2\Big(1-\frac{v_i^2}{c^2}\Big)^{-\frac{1}{2}}$ et $\vec{p}_i = m_i\vec{v}_i\Big(1-\frac{v_i^2}{c^2}\Big)^{-\frac{1}{2}}$ sont l'énergie et la quantité de mouvement de la particule. Effectuons une variation $\epsilon_t = \mathrm{d}T_1 = \mathrm{d}T_2 = -\mathrm{d}T_3 = -\mathrm{d}T_4 = -\mathrm{d}T_5$ de l'instant du choc et $\vec{\epsilon} = \mathrm{d}\vec{R}_1 = \mathrm{d}\vec{R}_2 = -\mathrm{d}\vec{R}_3 = -\mathrm{d}\vec{R}_4 = -\mathrm{d}\vec{R}_5$ de sa position ; alors : $\mathrm{d}\big(c^2\sum_i m_i\tau_i\big) = (E_1+E_2-E_3-E_4-E_5)\epsilon_t - (\vec{p}_1+\vec{p}_2-\vec{p}_3-\vec{p}_4-\vec{p}_5)\cdot\vec{\epsilon} = 0$. ϵ_t et $\vec{\epsilon}$ étant arbitraires ceci conduit à $\vec{p}_1+\vec{p}_2 = \vec{p}_3+\vec{p}_4+\vec{p}_5$ et $E_1+E_2 = E_3+E_4+E_5$. On peut montrer que l'extremum est un maximum. Ce principe (et des raisons d'homogénéité et de minimum) ont conduit à poser que l'action pour une particule libre est $S = -mc^2\tau$ et le lagrangien $L = -mc^2\Big(1-\frac{v^2}{c^2}\Big)^{\frac{1}{2}}$.

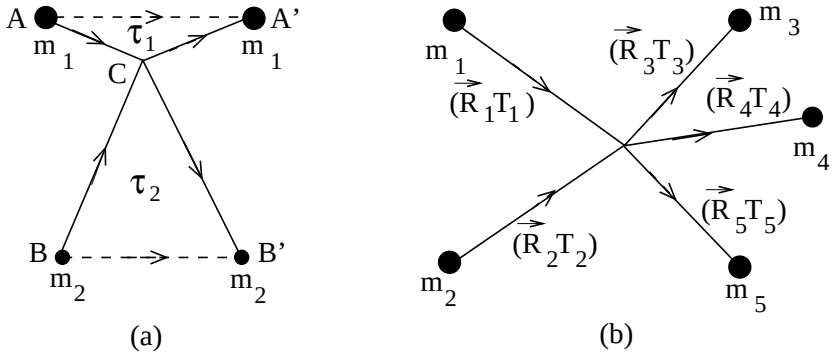


FIGURE 5

Approximation galiléenne. Rappelons (*cf.* section 4.3.3) que l'énergie de liaison, notée ici U, d'une particule composée (atome, noyau...) est la différence entre son énergie au repos mc^2 et la somme m^*c^2 des énergies de masse de ses constituants. En physique galiléenne $|U| \ll mc^2$, $v \ll c$, et m^* est conservée. L'approximation

$$-mc^2\Big(1-\frac{v^2}{c^2}\Big)^{\frac{1}{2}} \simeq -m^*c^2 + \frac{1}{2}mv^2 - U$$

montre que le lagrangien peut être pris égal à $L = \frac{1}{2}mv^2 - U$, ou à $\frac{1}{2}mv^2 - U(\overrightarrow{r},t)$ si $U(\overrightarrow{r},t)$ est l'énergie potentielle en présence d'un système extérieur. Si on considère des chocs, le principe devient "$\sum_i\Big(\frac{1}{2}m_i\frac{R_i^2}{T_i} - U_iT_i\Big)$ extremum" et on obtient les L.C. de la quantité de mouvement $m\vec{v}$ et de l'énergie $\frac{1}{2}m\vec{v}^{\,2} + U$.

9.2.2 Particules chargées et interactions électromagnétiques

▷ Principe de jauge et interactions électromagnétiques

Le principe d'extremum pour une particule libre s'écrit aussi

$$S = \int_{(A,t_1)}^{(B,t_2)}\Big(-mc^2\sqrt{1-\frac{v^2}{c^2}} - U + \vec{\mathcal{A}}\cdot\vec{v}\Big)\,\mathrm{d}t \quad \text{extremum}$$

quels que soient U et $\vec{\mathcal{A}}$ constants puisque les termes additionnels $\int -U\,\mathrm{d}t = U(t_1 - t_2)$ et $\int \vec{\mathcal{A}}\cdot\vec{v}\,\mathrm{d}t = \vec{\mathcal{A}}\cdot(\vec{r}_B - \vec{r}_A)$ ne dépendent que des bornes d'intégration. U et $\vec{\mathcal{A}}$ constituent des degrés de liberté supplémentaires. Le principe de jauge en physique consiste, lorsqu'une symétrie globale existe (ici la possibilité d'ajouter les termes contenant les constantes U et $\vec{\mathcal{A}}$ sans modifier les équations du mouvement), à introduire les interactions en imposant la **"localité" de la symétrie**, c'est-à-dire en faisant dépendre de $\vec{r}$ et de t les degrés de liberté liés à la symétrie. Alors que S avec U et $\vec{\mathcal{A}}$ constants (ou nuls) est l'action d'une particule libre, S avec $U(\vec{r},t) = qV(\vec{r},t)$ et $\vec{\mathcal{A}}(\vec{r},t) = q\vec{A}(\vec{r},t)$ où V et $\vec{A}$ sont les potentiels scalaire et vecteur est l'action pour une particule chargée en interaction avec le champ électromagnétique. Pour cette action il reste une symétrie de jauge associée au **changement de jauge** $V \to V - \dfrac{\partial\varphi}{\partial t}$ et $\vec{A} \to \vec{A} + \vec{\nabla}\varphi$ (avec $\varphi(\vec{r},t)$ arbitraire), changement qui revient à ajouter au lagrangien la dérivée totale $q\,\dfrac{\mathrm{d}\varphi}{\mathrm{d}t}$.

REMARQUE. En mécanique quantique l'**équation de Schrödinger** $i\hbar\partial_t\psi = \frac{1}{2m}\,(\frac{\hbar}{i}\,\overrightarrow{\nabla})^2\,\psi$ pour une particule libre possède la symétrie globale $\psi \to e^{i\varphi}\,\psi$ (φ constant). Pour rendre cette symétrie locale $\big(\varphi \to \varphi(\vec{r},t)\big)$ il faut effectuer dans l'équation les remplacements

$$i\hbar\partial_t \to i\hbar\partial_t - qV \quad \text{et} \quad \tfrac{\hbar}{i}\,\overrightarrow{\nabla} \to \tfrac{\hbar}{i}\,\overrightarrow{\nabla} - q\vec{A}\,,$$

qui introduisent l'interaction électromagnétique. L'équation obtenue est laissée invariante par le changement (symétrie de jauge) $V \to V - \partial_t\varphi$, $\vec{A} \to \vec{A} + \vec{\nabla}\varphi$ et $\psi \to \psi\,\exp\big(\frac{i}{\hbar}\,q\varphi\big)$.

▷ Mouvement d'une particule chargée dans un champ électromagnétique

Le lagrangien

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} + q(\vec{v}\cdot\vec{A} - V) \qquad \text{ou} \qquad L = \frac{1}{2}\,mv^2 + q(\vec{v}\cdot\vec{A} - V) \quad \text{pour } v \ll c$$

conduit à l'équation d'une particule soumise à la **force de Lorentz** :

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q(\vec{E} + \vec{v}\wedge\vec{B}) \quad \text{avec} \quad \vec{E} = -\partial_t\vec{A} - \overrightarrow{\mathrm{grad}}V \;,\; \vec{B} = \overrightarrow{\mathrm{rot}}\,\vec{A} \;,\; \vec{p} = m\vec{v}\,\big(1 - \frac{v^2}{c^2}\big)^{-\frac{1}{2}} \text{ (ou } m\vec{v}) \,.$$

DÉMONSTRATION : on écrit les équations $\dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial L}{\partial v_i}\right) = \dfrac{\partial L}{\partial x_i}$ avec, par exemple pour la première composante,

$$\frac{\partial L}{\partial x} = q\Big(-\frac{\partial V}{\partial x} + v_x\,\frac{\partial A_x}{\partial x} + v_y\,\frac{\partial A_y}{\partial x} + v_z\,\frac{\partial A_z}{\partial x}\Big) \quad , \quad \frac{\partial L}{\partial v_x} = p_x + qA_x \quad \text{et}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_x}\right) = \frac{\mathrm{d}p_x}{\mathrm{d}t} + q\Big(\frac{\partial A_x}{\partial t} + v_x\,\frac{\partial A_x}{\partial x} + v_y\,\frac{\partial A_x}{\partial y} + v_z\,\frac{\partial A_x}{\partial z}\Big) \,.$$

Attention. Il est important de remarquer que le **moment conjugué** $\vec{P}$ de $\vec{r}$ est différent de la quantité de mouvement $\vec{p}$ puisque $\vec{P} = \dfrac{\partial L}{\partial\vec{v}} = \vec{p} + q\vec{A}$. Le hamiltonien $H(\vec{P},\vec{r}) = \vec{P}\,.\,\vec{v} - L$ s'écrit (petit calcul) :

$$H(\vec{P},\vec{r}) = \sqrt{c^2\,(\vec{P} - q\vec{A})^2 + m^2c^4} + qV \quad \text{ou} \quad H(\vec{P},\vec{r}) = \frac{(\vec{P} - q\vec{A})^2}{2m} + qV \quad \text{pour } v \ll c \,.$$

Le fait que H s'exprime en fonction de $\vec{p}$, comme si $\vec{A}$, donc $\vec{B}$, n'existait pas, est trompeur. H doit être considéré, dans le formalisme de Hamilton, comme une fonction non de $\vec{p}$ mais de $\vec{P}$. Cette remarque, qui explique que les équations du mouvement ne sont pas les mêmes pour $\vec{B}=0$ et $\vec{B}\neq 0$, est encore plus importante quand on passe en physique quantique avec la correspondance $\vec{P}\leftrightarrow -i\hbar\overrightarrow{\nabla}$. Donnons trois exemples.

- **Particule non relativiste dans un champ $\vec{B}$ constant**. En utilisant l'expression $\vec{A}=\frac{1}{2}\,\vec{B}\wedge\vec{r}$ on obtient

$$H=\frac{(\vec{P}-q\vec{A})^2}{2m}+qV=\frac{\vec{P}^{\,2}}{2m}+qV-\vec{\mu}\cdot\vec{B}+\frac{q^2B^2r_\perp^2}{8m}\ ,$$

avec $r_\perp$ composante de $\vec{r}$ perpendiculaire à $\vec{B}$ et $\vec{\mu}=\dfrac{q}{2m}\,\vec{r}\wedge\vec{P}$. Considérant H comme un hamiltonien quantique, on voit qu'il diffère du cas $\vec{B}=0$ par deux termes. L'un $-\vec{\mu}\cdot\vec{B}$ est l'énergie d'un dipole magnétique $\vec{\mu}$; il est responsable du **paramagnétisme**. L'autre, présent même si $\vec{\mu}=0$, est l'énergie associée au **diamagnétisme**.

- **Courant supraconducteur**. Dans un supraconducteur les électrons de la surface de Fermi forment des états liés (paires de Cooper) pour lesquelles $<\vec{P}>=0$. Le courant électrique $\vec{j}=Nq<\vec{v}>$ vaut donc $\vec{j}_{\rm supra}=-N\frac{q^2}{m}\,\vec{A}$ et est lié à l'existence du potentiel vecteur $\vec{A}$. On peut montrer à partir de l'équation $\Delta\vec{A}=\frac{1}{\delta^2}\,\vec{A}$ conséquence de $\Delta\vec{A}=-\overrightarrow{\rm rot}\,\vec{B}=-\mu_0\vec{j}$ ($\delta=\frac{c}{\omega_p}$ avec $\omega_p^2=\frac{Nq^2}{m\epsilon_0}$) que le champ magnétique ne pénètre dans le matériau que sur une très faible épaisseur δ (**effet Meissner**).

- **Effet Aharonov-Bohm** (1956) (figure 6). Dans une expérience d'interférences de trous d'Young avec une source S d'électrons, la différence de phase en un point M du champ d'interférences s'écrit $\Delta\varphi=\frac{1}{\hbar}\oint_\gamma m\vec{v}\cdot{\rm d}\vec{r}$, où γ est le chemin fermé qui partant de S va à M en passant par un trou et revient à S en passant par l'autre trou. En présence d'un champ magnétique, $\Delta\varphi$ devient

$$\Delta\varphi=\tfrac{1}{\hbar}\oint_\gamma(m\vec{v}+q\vec{A})\cdot{\rm d}\vec{r}=2\pi\tfrac{\Delta l}{\lambda}+\tfrac{q\Phi}{\hbar}$$

où Δl est la différence des longueurs des deux chemins allant de S à M, $\lambda=\frac{h}{mv}$ la longueur d'onde de de Broglie et Φ le flux de $\vec{B}$ à travers γ. La présence du champ se traduit donc par un déplacement de la figure d'interférences d'un nombre $\frac{q\Phi}{h}$ d'interfranges (effet Aharonov-Bohm). Cet effet a lieu même si les électrons ne passent que dans des régions où $\vec{B}=0$.

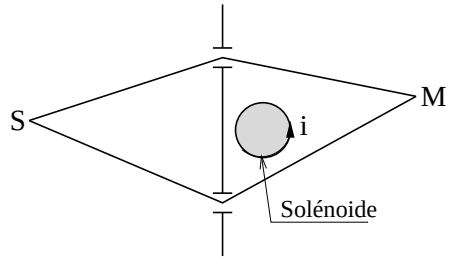


FIGURE 6

9.2.3 Temps propre et gravitation ; courbure de l'espace-temps

▷ Gravitation et géométrie ; équation des géodésiques

Dans un potentiel gravitationnel newtonien $V(\vec{r})$, dont dérive le champ $\vec{g} = -\overrightarrow{\text{grad}}\, V$, l'énergie potentielle d'une particule de masse m est $U(\vec{r}) = mV(\vec{r})$ (**identité des masses gravitationnelle et inertielle**). Si $\dfrac{V}{c^2} \ll 1$ et $\dfrac{v^2}{c^2} \ll 1$ le lagrangien galiléen $\dfrac{1}{2}\,mv^2 - U$ équivaut à $L \simeq -mc^2\Big(1-\dfrac{v^2}{c^2}+\dfrac{2V}{c^2}\Big)^{\frac{1}{2}}$. Ceci revient à introduire un temps propre "modifié" (on suppose $V = 0$ à l'infini) :

$$\tau = \int_{(A,t_1)}^{(B,t_2)} \mathrm{d}\tau \quad \text{avec} \;\; \mathrm{d}\tau^2 = (1 + \frac{2V}{c^2})\,\mathrm{d}t^2 - \frac{\mathrm{d}\vec{r}^2}{c^2}\ .$$

Les mouvements de chute libre, qui extrémalisent τ, apparaissent alors comme des **géodésiques** de l'espace-temps muni d'une **métrique** $c^2\mathrm{d}\tau^2$ déterminée par la gravitation (analogie avec les géodésiques de l'espace qui minimisent les distances). Ce point de vue géométrique sur la gravitation est renforcé par le fait que, pour un mouvement quelconque, τ est effectivement ce que mesure toute horloge interne à un système physique (par exemple des horloges atomiques).

Dans un système arbitraire de coordonnées x^μ ($\mu = 0, 1, 2, 3$) le temps propre s'écrit a priori sous la forme

$$\tau = \int \mathrm{d}\tau \quad \text{avec} \quad c^2\mathrm{d}\tau^2 = g_{\mu\nu}(x)\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \;>\; 0 \quad (g_{\mu\nu} = g_{\nu\mu})\ .$$

(Dans toute cette section la sommation sur un indice est sous-entendue si celui-ci figure une fois en position haute et une fois en position basse ; dans le cas newtonien ci-dessus : $x^0 = ct$, $\{x^i\} \equiv \vec{r}$ ($i = 1, 2, 3$), $g_{00} = 1 + 2Vc^{-2}$, $g_{ii} = -1$ et $g_{\mu\nu} = 0$ si $\mu \neq \nu$). Les géodésiques $x^\mu(\lambda)$, paramétrées par λ, obéissent alors aux E.D. d'Euler-Lagrange relatives au lagrangien

$$L = \frac{c\,\mathrm{d}\tau}{\mathrm{d}\lambda} = \Big[g_{\mu\nu}\,\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\,\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\Big]^{\frac{1}{2}}\ .$$

Pour la coordonnée x^α on obtient (calcul analogue à celui de la section 9.1.1 pour l'obtention de la loi de Descartes à partir du principe de Fermat)

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big[\frac{1}{2}\,\frac{\mathrm{d}\lambda}{\mathrm{d}\tau}\,2g_{\mu\alpha}\,\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\Big] = \frac{1}{2}\,\frac{\mathrm{d}\lambda}{\mathrm{d}\tau}\,(\partial_\alpha\, g_{\mu\nu})\,\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\,\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \qquad \Big(\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}\Big)\ ,$$

ou, en introduisant la **quadrivitesse** $u^\mu = \dfrac{\mathrm{d}x^\mu}{c\,\mathrm{d}\tau}$ (analogue du vecteur unitaire $\hat{u}$) :

$$\frac{\mathrm{d}}{c\,\mathrm{d}\tau}\,\big[g_{\mu\alpha}\,u^\mu\big] = \frac{1}{2}\,(\partial_\alpha\, g_{\mu\nu})\,u^\mu\,u^\nu\ .$$

Dans le cas newtonien (avec $V \ll c^2$ et $v^2 \ll c^2$), on retrouve la conservation de l'énergie $\frac{1}{2}v^2 + V$ pour $\alpha = 0$ et la loi de Newton $\vec{a} = -\overrightarrow{\text{grad}}\, V$ pour $\alpha = 1, 2, 3$.

REMARQUES. 1) En présence d'un champ de gravitation, il n'y a pas de forme "canonique" pour la métrique. Seules des hypothèses (stationnarité, isotropie, comportement à l'infini... *cf.* la métrique de

Schwarzschild) conduisent à une écriture plus simple pour un choix de coordonnées. Un **changement de coordonnées** arbitraire $\{x^\mu\} \to \{x'^\alpha\}$ conduisant à récrire l'invariant $c^2\,\mathrm{d}\tau^2$ sous la forme

$$c^2\,\mathrm{d}\tau^2 = g'_{\alpha\beta}(x')\,\mathrm{d}x'^\alpha\,\mathrm{d}x'^\beta \quad \text{avec} \quad g'_{\alpha\beta}(x') = g_{\mu\nu}(x)\,\tfrac{\partial x^\mu}{\partial x'^\alpha}\,\tfrac{\partial x^\nu}{\partial x'^\beta}\ ,$$

il n'existe évidemment pas de choix des quatre fonctions x^μ qui permette de rendre par exemple les dix fonctions $g'_{\alpha\beta}$ constantes comme pour un espace-temps plat.

2) En mécanique newtonienne la relation $\overrightarrow{a} = \overrightarrow{g}$ peut faire penser qu'on peut éliminer la gravitation en se plaçant dans un référentiel en chute libre, mais c'est inexact car $\overrightarrow{g}$ n'est jamais rigoureusement uniforme. Dans ce référentiel, un point $\overrightarrow{r}$ en chute libre voisin de l'origine vérifie $\ddot{\overrightarrow{r}} = \overrightarrow{g}(\vec{r}) - \overrightarrow{g}(0) = \delta\vec{g}(\vec{r}) \neq 0$; le champ $\delta\vec{g}(\vec{r})$, appelé **champ de marée** car ses équipotentielles correspondent à des surfaces de niveau de masses fluides en chute libre, dépend des dérivées secondes de $V(\vec{r})$ en $\overrightarrow{r} = 0$ et ne peut être annulé. De même, on montre que le changement $\{x^\mu\} \to \{x'^\alpha\}$ permet, au mieux, de mettre en un point $\mathrm{d}\tau^2$ sous une forme diagonale (supposée être $\mathrm{d}t^2 - c^{-2}\mathrm{d}\overrightarrow{r}^2$ pour garder la **symétrie de Lorentz**) et d'annuler en ce point les dérivées premières des $g'_{\alpha\beta}$ (mais pas les dérivées secondes si l'espace-temps est courbe).

▷ Tests de la métrique de Schwarzschild

Dans le champ dû à une masse M sphérique on a exactement (résultat admis ici) :

$$\mathrm{d}\tau^2 = \Big(1 - \frac{2GM}{rc^2}\Big)\,\mathrm{d}t^2 - \frac{r^2}{c^2}\,(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2) - \frac{\mathrm{d}r^2}{c^2\Big(1 - \frac{2GM}{rc^2}\Big)}\ .$$

- **Redshift gravitationnel**. Si une source émet des "tops" en $\vec{r}_1$ fixe, à intervalles de temps $\Delta t_1 = T$ réguliers, ils sont reçus en $\vec{r}_2$ fixe, à intervalles de temps réguliers eux aussi $\Delta t_2 = T$ (car la métrique et donc les temps de propagation ne dépendent pas de t) ; mais les fréquences $\nu_1 = (\Delta\tau_1)^{-1}$ et $\nu_2 = (\Delta\tau_2)^{-1}$ mesurées en ces points par des horloges "physiques" sont différentes : $\nu_2/\nu_1 = \big(1 - \frac{2GM}{r_1c^2}\big)^{\frac12}/\big(1 - \frac{2GM}{r_2c^2}\big)^{\frac12}$ (**redshift** $\nu_2 < \nu_1$ si $r_2 > r_1$).

- **Trajectoires d'un point matériel**. Si on paramètre un mouvement par $t(\lambda)$ et $\vec{r}(\lambda)$, on a $d\tau = L\,d\lambda$ avec

$$L = \left(\Big(1 - \frac{2\alpha}{r}\Big)\,t'^2 - \Big(1 - \frac{2\alpha}{r}\Big)^{-1} r'^2 - r^2\theta'^2 - r^2\,\sin^2\theta\,\varphi'^2\right)^{\frac12} \qquad (\alpha = GM \ \text{ et } \ c = 1)\ ,$$

f' désignant la dérivée de f par rapport à λ. Un argument simple montrant que la trajectoire est plane consiste à remarquer que $\theta = \dfrac{\pi}{2} = \mathrm{C^{ste}}$ correspond à un extremum de l'action, puisqu'en posant $\theta = \dfrac{\pi}{2} + \epsilon$ la variation de $\int L\,\mathrm{d}\lambda$ est en ϵ^2. De plus comme les variables t et φ n'apparaissent pas dans L, les quantités $\dfrac{\partial L}{\partial t'}$ et $\dfrac{\partial L}{\partial \varphi'}$ sont constantes. Si on note $\dot{f}$ la dérivée de f par rapport à τ on a

$$\Big(1 - \frac{2\alpha}{r}\Big)\,\dot{t} = A \quad , \quad r^2\,\dot{\varphi} = C \quad \text{et} \quad \Big(1 - \frac{2\alpha}{r}\Big)\,\dot{t}^2 - \Big(1 - \frac{2\alpha}{r}\Big)^{-1}\dot{r}^2 - r^2\,\dot{\varphi}^2 = \frac{\mathrm{d}\tau^2}{\mathrm{d}\tau^2} = 1\ ,$$

d'où $A^2 - \dot{r}^2 - C^2r^{-2}\Big(1 - \dfrac{2\alpha}{r}\Big) = 1 - \dfrac{2\alpha}{r}$. En posant $u = \dfrac{1}{r}$ et en utilisant $\dfrac{\mathrm{d}}{\mathrm{d}\tau} = \dot{\varphi}\,\dfrac{\mathrm{d}}{\mathrm{d}\varphi} = \dfrac{C}{r^2}\,\dfrac{\mathrm{d}}{\mathrm{d}\varphi}$ on obtient, après dérivation par rapport à φ, l'équation de la trajectoire :

$$\frac{\mathrm{d}^2u}{\mathrm{d}\varphi^2} + u - \frac{\alpha}{C^2} - 3\alpha u^2 = 0\ .$$

- **Déplacement du périhélie**. Dans le problème de Kepler habituel (*cf.* section 2.3.3) le terme non linéaire en u^2 est absent et la solution est l'ellipse $u_0(\varphi) = \frac{1}{p}\,(1 + e\,\cos\,\varphi)$ avec $p = \frac{C^2}{\alpha}$. Si on considère ce terme comme une perturbation, en posant $u(\varphi) = u_0(\varphi) + \delta u(\varphi)$ et en ne prenant en compte dans $3\alpha u_0^2$ que le terme résonnant (*cf.* section 6.5.1), un calcul élémentaire conduit à $\delta u(\varphi) = 3\alpha p^{-2} e\varphi\,\sin\,\varphi$. Tout se passe donc comme si dans $u_0(\varphi)$ il fallait remplacer cos φ par $\cos\,(1 - \frac{3\alpha}{p})\varphi$. Ainsi r ne reprend la même valeur que lorsque φ varie de $2\pi + \Delta\varphi$ avec $\Delta\varphi = \frac{6\pi\alpha}{p} = \frac{6\pi\alpha}{a(1-e^2)}$.

- **Déviation des rayons lumineux**. Pour étudier la déviation des rayons lumineux (photons de masse nulle) il faut considérer la limite $\mathrm{d}\tau = 0$ ce qui correspond à C infini dans l'équation pour u. Si on néglige le terme en u^2, la solution est la droite $u_0(\varphi) = \frac{1}{R}\,\cos\,\varphi$ parallèle à l'axe y et passant à une distance R de O (pas de déviation du photon) ; la prise en compte de ce terme dans un calcul de perturbation donne $\delta u(\varphi) = \frac{3\alpha}{2R^2}\,\big(1 - \frac{\cos\,2\varphi}{3}\big)$. En écrivant qu'à l'infini φ est tel que $u(\varphi) = 0$ et sachant qu'alors $\varphi = \pm(\frac{\pi}{2} + \epsilon)$, on obtient la déviation des rayons lumineux $\delta\varphi = 2\epsilon = \frac{4\alpha}{R}$.

▷ Transport parallèle et courbure

Le premier membre de l'équation des géodésiques s'écrit aussi

$$\tfrac{\mathrm{d}}{c\,\mathrm{d}\tau}\big(g_{\mu\alpha}\,u^\mu\big) = g_{\mu\alpha}\,\tfrac{\mathrm{d}u^\mu}{c\,\mathrm{d}\tau} + \big(\partial_\nu g_{\mu\alpha}\big)\,u^\mu u^\nu = g_{\mu\alpha}\,\tfrac{\mathrm{d}u^\mu}{c\,\mathrm{d}\tau} + \tfrac{1}{2}\big(\partial_\nu g_{\mu\alpha} + \partial_\mu g_{\nu\alpha}\big)\,u^\mu u^\nu\ .$$

En multipliant les deux membres de cette équation par les nombres $g^{\beta\alpha} = g^{\alpha\beta}$ tels que $g^{\beta\alpha}g_{\mu\alpha} = 1$ si $\beta = \mu$ et égal à 0 si $\beta \neq \mu$ (nombres qui constituent les éléments de la matrice symétrique inverse de celle ayant pour éléments $\{g_{\mu\alpha}\}$) on obtient (après sommation sur α) :

$$\tfrac{\mathrm{d}u^\beta}{c\,\mathrm{d}\tau} = -\Gamma^\beta_{\mu\nu}\,u^\mu u^\nu \quad \text{où} \quad \Gamma^\beta_{\mu\nu} = g^{\beta\alpha}\tfrac{1}{2}(\partial_\nu g_{\mu\alpha} + \partial_\mu g_{\nu\alpha} - \partial_\alpha g_{\mu\nu}) = \Gamma^\beta_{\nu\mu} \quad (\textbf{symboles de Christoffel}).$$

Ceci implique que, dans un déplacement infinitésimal le long de la géodésique, les composantes de la quadrivitesse varient comme

$$\mathrm{d}u^\beta = -\Gamma^\beta_{\mu\nu}\mathrm{d}x^\mu u^\nu\ .$$

L'intérêt de cette relation est qu'elle correspond au **transport parallèle** de la quadrivitesse. En effet, comme pour la sphère (*cf.* section 3.2.6), le transport de x en $x + \mathrm{d}x$ d'un quadrivecteur est défini de sorte qu'il conserve sa norme et sa projection sur la géodésique. Or u^μ est un **quadrivecteur** (quantité, qui en x, se transforme comme $\mathrm{d}x^\mu$ dans un changement de coordonnées), il est tangent à la géodésique en tout point (car proportionnel à $\mathrm{d}x^\mu$) et est unitaire ($u^\mu(x)\,g_{\mu\nu}(x)\,u^\nu(x) = \mathrm{d}\tau^2/\mathrm{d}\tau^2 = 1$). On peut vérifier que pour un quadrivecteur quelconque A^μ (non nécessairement tangent aux géodésiques et non unitaire) la même relation

$$\delta A^\beta = -\Gamma^\beta_{\mu\nu}\mathrm{d}x^\mu\,A^\nu$$

convient pour définir son transport parallèle car elle assure la conservation du produit scalaire : $(A^\mu + \delta A^\mu)\,g_{\mu\nu}(x + \mathrm{d}x)\,(B^\nu + \delta B^\nu) = A^\mu\,g_{\mu\nu}(x)\,B^\nu$. (Pour cela écrire $g_{\mu\nu}(x + \mathrm{d}x) = g_{\mu\nu}(x) + (\partial_\rho g_{\mu\nu})\,\mathrm{d}x^\rho$ et utiliser la définition des Γ.)

La **courbure** est définie, comme pour la sphère, à partir du transport parallèle d'un quadrivecteur le long d'un circuit fermé orienté. Pour un parallélogramme infinitésimal situé en x de côtés $\delta_1 x^\mu$, $\delta_2 x^\mu$, $-\delta_1 x^\mu$, $-\delta_1 x^\mu$, et dont les projections des aires sur les plans (x^μ, x^ν) sont $\delta S^{\mu\nu} = \delta_1 x^\mu\,\delta_2 x^\nu - \delta_1 x^\nu\,\delta_2 x^\mu$, la variation d'un quadrivecteur dans le transport s'écrit (généralisation de la formule de Stokes de la section 7.2.2)

$$\Delta A^\beta = -\oint \Gamma^\beta_{\nu\alpha}A^\alpha\,\mathrm{d}x^\nu = -\tfrac{1}{2}[\partial_\mu(\Gamma^\beta_{\nu\alpha}A^\alpha) - \partial_\nu(\Gamma^\beta_{\mu\alpha}A^\alpha)]\,\delta S^{\mu\nu}\ ,$$

ou encore ($\partial_\mu A^\alpha$ au point x étant déduit de δA^α, réflexion laissée au lecteur)

$$\Delta A^\beta = -\tfrac{1}{2}R^\beta_{\alpha\mu\nu}\,A^\alpha\,\delta S^{\mu\nu} \quad \text{avec} \quad R^\beta_{\alpha\mu\nu} = (\partial_\mu\Gamma^\beta_{\alpha\nu} - \partial_\nu\Gamma^\beta_{\alpha\mu}) + (\Gamma^\beta_{\sigma\mu}\Gamma^\sigma_{\alpha\nu} - \Gamma^\beta_{\sigma\nu}\Gamma^\sigma_{\alpha\mu})\ .$$

Les coefficients $R^\beta_{\alpha\mu\nu}$, fonctions du premier degré des dérivées secondes des $g_{\mu\nu}$, caractérisent la **courbure de Riemann** de l'espace-temps. (Rappelons que pour la sphère la rotation d'un vecteur lors de son transport est aussi proportionnelle à la surface du circuit et à la courbure.) Les grandeurs dérivées

$$R_{\alpha\mu} = R^\beta_{\alpha\mu\beta}\ \ (\textbf{courbure de Ricci}) \quad , \quad R = g^{\alpha\mu}R_{\alpha\mu}\ \ (\textbf{courbure scalaire})\ ,$$

qui ont les propriétés remarquables respectivement, de se transformer comme les $g_{\alpha\mu}$ (avec $R_{\alpha\mu} = R_{\mu\alpha}$), et d'être invariante, jouent un rôle important pour les équations d'Einstein de la gravitation (*cf.* section 9.3.3).

JUSTIFICATION des propriétés de transformation de $R_{\mu\nu}$ et de R dans un changement de coordonnées. Elles résultent des deux relations (dont la démonstration est donnée ci-après) :

$$\text{(a)}\quad R'^{\delta}_{\gamma\rho\sigma} = \frac{\partial x'^{\delta}}{\partial x^{\beta}}\frac{\partial x^{\alpha}}{\partial x'^{\gamma}}\frac{\partial x^{\mu}}{\partial x'^{\rho}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\,R^{\beta}_{\alpha\mu\nu} \quad \text{et} \quad \text{(b)}\quad \mathrm{g}'^{\gamma\delta} = \frac{\partial \mathrm{x}'^{\gamma}}{\partial \mathrm{x}^{\mu}}\frac{\partial \mathrm{x}'^{\delta}}{\partial \mathrm{x}^{\nu}}\,\mathrm{g}^{\mu\nu}\ .$$

Pour $R_{\mu\nu}$, il suffit de faire $\delta=\sigma$ dans (a) en remarquant que $\frac{\partial x^{\nu}}{\partial x'^{\sigma}}\frac{\partial x'^{\sigma}}{\partial x^{\beta}}=1$ si $\beta=\nu$ et 0 si $\beta\neq\nu$. Pour R, on déduit de (b) l'égalité : $g'^{\gamma\delta}R'_{\gamma\delta} = \frac{\partial x'^{\gamma}}{\partial x^{\mu}}\frac{\partial x'^{\delta}}{\partial x^{\nu}}\,g^{\mu\nu}\,\frac{\partial x^{\rho}}{\partial x'^{\gamma}}\frac{\partial x^{\sigma}}{\partial x'^{\delta}}\,R_{\rho\sigma} = g^{\mu\nu}\,R_{\mu\nu}$. La relation (a) s'obtient à partir de l'expression définissant la courbure de Riemann $R^{\beta}_{\alpha\mu\nu}$ en exprimant les quadivecteurs ΔA^{β}, A^{α}, δx^{μ} et δx^{ν} en x en fonction des mêmes quantités primées, par exemple $\delta x^{\nu} = \frac{\partial x^{\nu}}{\partial x'^{\sigma}}\,\delta x'^{\sigma}$. La relation (b) se vérifie en observant que $g'^{\gamma\alpha}g'_{\alpha\beta}=1$ si $\beta=\gamma$ et 0 si $\beta\neq\gamma$. Enfin, la démonstration de $R_{\alpha\mu}=R_{\mu\alpha}$ revient à vérifier que $\partial_{\mu}\Gamma^{\beta}_{\beta\alpha}$ est symétrique en μ et α ; ceci est vrai car $\Gamma^{\beta}_{\beta\alpha}=\frac{1}{2}g^{\mu\nu}\,\partial_{\alpha}\,g_{\mu\nu}$ (*cf.* définition des Γ) s'écrit aussi $\frac{1}{2}\,\partial_{\alpha}\ln\Delta$ avec $\Delta=|\det g_{\mu\nu}|$. (En effet, pour tout matrice $\mathbf{A}$: $\delta(\det\mathbf{A}) = \mathrm{tr}(\tilde{\mathbf{A}}^{t}\,\delta\mathbf{A}) = (\det\mathbf{A})\,\mathrm{tr}(\mathbf{A}^{-1}\,\delta\mathbf{A})$ avec $\tilde{\mathbf{A}}$ matrice des cofacteurs *cf.*section 4.2.2.)

9.3 CHAMPS ET PRINCIPES DE MOINDRE ACTION

9.3.1 Equations de Maxwell ; action de Schwarzschild

On obtient les deux équations de Maxwell impliquant les sources en rendant extremum, par rapport aux potentiels $\overrightarrow{A}$ et V, l'**action invariante**

$$S = \iiiint \big(\frac{1}{2}\epsilon_0(\overrightarrow{E}^{\,2} - c^2\overrightarrow{B}^{\,2}) + (\overrightarrow{j}\cdot\overrightarrow{A} - \rho V)\big)\,\mathrm{d}^3r\,\mathrm{d}t\ ,$$

dans laquelle $\overrightarrow{E} = -\partial_t\overrightarrow{A} - \overrightarrow{\mathrm{grad}}V$, $\overrightarrow{B} = \overrightarrow{\mathrm{rot}}\,\overrightarrow{A}$ et le terme d'interaction avec les charges généralise le terme $q\,(\overrightarrow{v}\cdot\overrightarrow{A} - V)$introduit à la section 9.2.2.
DÉMONSTRATION. La variation de $\overrightarrow{A}$ se traduit sur S par :

$$\delta S = \iiiint \big(\epsilon_0\,\overrightarrow{E}\cdot(-\partial_t\delta\vec{A}) - c^2\overrightarrow{B}\cdot\overrightarrow{\mathrm{rot}}\,\delta\vec{A} + \overrightarrow{j}\cdot\delta\vec{A}\big)\,\mathrm{d}^3r\,\mathrm{d}t\ .$$

Après intégration par parties ($\delta\vec{A}$ étant pris nul à l'infini), et utilisation de $\mathrm{div}(\delta\vec{A}\wedge\overrightarrow{B}) = \overrightarrow{B}\cdot\overrightarrow{\mathrm{rot}}\,\delta\vec{A} - \delta\vec{A}\cdot\overrightarrow{\mathrm{rot}}\,\overrightarrow{B}$ (et $\epsilon_0\mu_0c^2=1$), δS s'écrit :

$$\delta S = \iiiint \big(\epsilon_0\partial_t\overrightarrow{E} - \frac{1}{\mu_0}\overrightarrow{\mathrm{rot}}\,\overrightarrow{B} + \overrightarrow{j}\big)\cdot\delta\vec{A}\,\mathrm{d}^3r\,\mathrm{d}t\ .$$

La condition $\delta S=0$ quelle que soit la fonction $\delta\vec{A}(\overrightarrow{r},t)$ conduit à $\overrightarrow{\mathrm{rot}}\,\overrightarrow{B} = \mu_0\,(\overrightarrow{j} + \epsilon_0\,\partial_t\overrightarrow{E})$. On obtient de même $\mathrm{div}\,\overrightarrow{E} = \frac{\rho}{\epsilon_0}$ en faisant varier V (et en utilisant $\overrightarrow{E}\cdot(-\overrightarrow{\mathrm{grad}}\delta V) = -\mathrm{div}(\overrightarrow{E}\,\delta V) + \delta V\,\mathrm{div}\,\overrightarrow{E}$). L'invariance de S résulte de ce que dans une transformation de Lorentz (*cf.* section 4.3.3) $\overrightarrow{E}^{\,2} - c^2\overrightarrow{B}^{\,2}$ est invariant de même que $\rho V - \overrightarrow{j}\cdot\overrightarrow{A}$ (produit scalaire de deux quadrivecteurs) et $\mathrm{d}^3r\,\mathrm{d}t$ (car le jacobien des transformations vaut à 1). Elle garantit l'**invariance relativiste** des équations de Maxwell. Le lecteur vérifiera que l'autre invariant $\epsilon_0 c\,\overrightarrow{E}\cdot\overrightarrow{B}$ ne contribue pas à δS.

9.3.2 Equation de Schrödinger ; théorème adiabatique

▷ Equation de Schrödinger

L'équation $i\hbar\partial_t\underline{\psi} = H\underline{\psi}$, qui est une E.D.P. si le vecteur d'état $\underline{\psi}(t)$ est une fonction d'onde $\underline{\psi}(\overrightarrow{r},t)$, entraîne :

$$S = \int_{t_1}^{t_2} (\underline{\psi}\,,\, (i\hbar\partial_t - H)\underline{\psi})\,\mathrm{d}t \quad \text{extremum .}$$

VÉRIFICATION : $\delta S = \int_{t_1}^{t_2}(\delta\underline{\psi}\,,\,(i\hbar\partial_t - H)\underline{\psi})\,\mathrm{d}t + \int_{t_1}^{t_2}(\underline{\psi}\,,\,(i\hbar\partial_t - H)\delta\underline{\psi})\,\mathrm{d}t$. Après intégration par parties et compte tenu de l'hermiticité de H, le deuxième terme s'écrit aussi $\int_{t_1}^{t_2}((i\hbar\partial_t - H)\underline{\psi}\,,\,\delta\underline{\psi})\,\mathrm{d}t$. La réciproque , facile à montrer, est laissée au lecteur.

Etat stationnaire (H indépendant du temps, $\underline{\psi}(t) = e^{-i\frac{E}{\hbar}t}\,\underline{\psi}(0)$ vecteur propre de H). Le principe conduit à $(\underline{\psi}\,,\,H\underline{\psi})$ extremum (*cf.* section 4.2.3, calculs de perturbation).

▷ Théorème adiabatique quantique

. Soient $\underline{\psi}_n(t)$ et $E_n(t)$ les vecteurs propres (base orthonormée) et valeurs propres "instantanés" d'un hamiltonien $H(t)$ qui dépend lentement du temps, et dont on suppose le spectre discret. Si l'échelle de temps de variation T_H de $H(t)$ est grande devant les périodes $T_{nm} = \dfrac{\hbar}{E_m - E_n}$ (hypothèse adiabatique) on a, pour un état quelconque $\underline{\psi}(t) = \sum_n \rho_n(t)\, e^{i\varphi_n(t)}\,\underline{\psi}_n(t)$ (avec ρ_n réel) :

$$\rho_n = \text{ constante} \qquad \text{et} \qquad \dot{\varphi}_n = -\frac{E_n(t)}{\hbar} + i\,(\underline{\psi}_n(t)\,,\,\dot{\underline{\psi}}_n(t))\ .$$

DÉMONSTRATION : $S = \int \sum_n \rho_n \left(i\hbar\dot{\rho}_n - \hbar\dot{\varphi}_n\rho_n - E_n\rho_n + \sum_m i\hbar\rho_m e^{i(\varphi_m-\varphi_n)}(\underline{\psi}_n\,,\,\dot{\underline{\psi}}_m)\right)\mathrm{d}t$. En négligeant les termes en $e^{i(\varphi_m-\varphi_n)}$ $(n \neq m)$ qui oscillent "rapidement" et contribuent peu à S (approximation adiabatique), et en oubliant le terme $\sum_n \rho_n\dot{\rho}_n$ (dérivée totale), le lagrangien s'écrit :

$$\mathcal{L} = \sum_n \rho_n^2 \left(-\hbar\dot{\varphi}_n - E_n + i\hbar(\underline{\psi}_n(t)\,,\,\dot{\underline{\psi}}_n(t))\right)\ .$$

Les E.D. d'Euler-Lagrange pour les variables ρ_n et φ_n donnent le résultat annoncé.

Conséquence : puisque les ρ_n sont constants, une évolution adiabatique n'induit pas de transition entre niveaux d'énergie ; en particulier si l'état initial est l'état propre instantané $\underline{\psi}_n(0)$, l'état à l'instant t sera, à une phase près, l'état propre instantané $\underline{\psi}_n(t)$. Cela conduit en physique statistique des systèmes en équilibre thermique à une distinction microscopique entre travail et échanges thermiques réversibles :

$$\mathrm{d}(U = \sum N_n\,E_n) = \sum_n N_n\,\mathrm{d}E_n + \sum_n E_n\,\mathrm{d}N_n = \delta W_{rev} + \delta Q_{rev}\ .$$

Dans une transformation adiabatique les populations N_n des niveaux d'énergie sont constantes et on retrouve la définition thermodynamique de l'adiabaticité $\delta Q_{rev} = 0$. De façon générale un **travail réversible** modifie les niveaux d'énergie tandis qu'un **apport thermique** modifie leurs populations. Par exemple pour une particule libre dans une boite cubique de côté a, on a $E_{npq} = \dfrac{\hbar^2}{2m}\,\dfrac{\pi^2}{a^2}\,(n^2+p^2+q^2)$ et $\dfrac{\delta E_{npq}}{E_{npq}} = -2\,\dfrac{\mathrm{d}a}{a} = -\dfrac{2}{3}\,\dfrac{\mathrm{d}V}{V}$.

En écrivant le travail $\delta W_{rev} = \sum N_{npq}\,\delta E_{npq} = -\dfrac{2}{3}\,U\,\dfrac{\mathrm{d}V}{V} = -P\,\mathrm{d}V$, on retrouve l'expression bien connue $P = \dfrac{2}{3}\,\dfrac{U}{V}$ de la **pression cinétique** d'un gaz parfait monoatomique.

Phase de Berry. On appelle ainsi le déphasage supplémentaire $\varphi_B = \int_{t_1}^{t_2} i(\underline{\psi}_n(t)\,,\,\dot{\underline{\psi}}_n(t))\,\mathrm{d}t$ lié au fait que le système évolue dans un environnement lentement variable. Il est mis en évidence dans des expériences d'interférences où, sur les deux chemins, les $E_n(t)$ sont les mêmes mais les $\psi_n(t)$ sont différents (spins $\frac{1}{2}$ se déplaçant "lentement" dans un champ magnétique $\vec{B}(\vec{r})$ dont seule la direction varie).

Transport quantique. La phase de Berry correspond à un transport naturel entre deux états quantiques voisins ψ et $\psi + \delta\psi$. En effet la minimisation de $\parallel (\psi + \delta\psi)e^{i\delta\varphi} - \psi \parallel^2$, rendant ces deux états les plus proches possible, conduit à l'ajustement de phase $\delta\varphi = i(\psi, \delta\psi)$.

9.3.3 Equations d'Einstein de la gravitation ; action de Hilbert

Les équations d'Einstein relient la courbure de Ricci de l'espace-temps à la matière, décrite de façon continue par $T_{\mu\nu}$ (voir interprétation physique ci-dessous) :

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = -\tfrac{8\pi G}{c^4}T_{\mu\nu} \quad (T_{\mu\nu} = T_{\nu\mu})\ .$$

Ce sont des E.D.P. du second ordre pour les coefficients de la métrique. Elles équivalent dans le vide ($T_{\mu\nu} = 0$) à $R_{\mu\nu} = 0$ (car $g^{\mu\nu}(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R) = R - 2R = 0$), et dans la matière à $R_{\mu\nu} = -\frac{8\pi G}{c^4}(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T)$ avec $T = g^{\alpha\beta}T_{\alpha\beta}$. Dans le cas newtonien $\mathrm{d}\tau^2 = (1+2Vc^{-2})\,\mathrm{d}t^2 - c^{-2}\,\mathrm{d}\vec{r}^2$ où les seules dérivées non nulles des $g_{\mu\nu}$ sont les $\partial_i g_{00}$ ($i = 1, 2, 3$) et où on ne garde que les termes proportionnels à V ou ses dérivées, on peut vérifier que les $R_{\mu\nu}$ se réduisent à $R_{00} \simeq -\Delta V/c^2$; comme dans cette limite ne reste aussi que $T_{00} \simeq \rho c^2$ (ρ masse volumique), on retrouve l'E.D.P. de **Poisson**

$$\Delta V = 4\pi G\rho\ .$$

(Pour une masse ponctuelle à l'origine cette E.D.P a pour solution $V = -Gm/r$, de même que pour une charge électrique ponctuelle $V = q/(4\pi\epsilon_0 r)$ correspond à $\Delta V = -\rho/\epsilon_0$.)

Pour un faisceau monocinétique de particules de masse m et de densité n_0 dans le référentiel local où elles sont au repos : $T_{\mu\nu} = g_{\alpha\mu}g_{\beta\nu}T^{\alpha\beta}$ avec $T^{\alpha\beta} = n_0 mc^2 u^\alpha u^\beta$ ($u^\alpha = \mathrm{d}x^\alpha/c\,\mathrm{d}\tau$ est la quadrivitesse introduite à la section 9.2.3 et les coefficients $g_{\alpha\mu}$ et $g_{\beta\nu}$ assurent que $T_{\mu\nu}$ se transforme comme $R_{\mu\nu}$ et $g_{\mu\nu}$). Dans un espace-temps plat, et dans un référentiel où $\vec{v}$ est la vitesse des particules et $n = \gamma n_0$ ($\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$) est leur densité on a :

$$T^{00} = n\gamma mc^2 \quad , \quad \{T^{0i}\} \equiv n\gamma mc\,\vec{v} \quad , \quad \mathrm{T}^{11} = n\gamma m v_x^2 \cdots \quad , \quad \mathrm{T}^{12} = n\gamma m v_x v_y \cdots\ .$$

T^{00} est donc la **densité d'énergie**, les cT^{0i} correspondent à la **densité de courant d'énergie** (et les $c^{-1}T^{0i}$ à la **densité de quantité de mouvement**) et les T^{ij} décrivent un flux de quantité de mouvement (*cf.* l'analogue non relativiste section 3.2.4). Ces interprétations physiques sont pour l'essentiel les mêmes en espace-temps courbe, que les $T^{\mu\nu}$ soient associés à de la matière ou à des champs (comme le champ électromagnétique).

On montre aussi (*cf.* ouvrages spécialisés) que les équations d'Einstein obéissent à un principe variationnel, les variations portant ici sur les $g_{\mu\nu}$, dont l'**action invariante** par changement de coordonnées est de la forme (action d'**Hilbert**) :

$$S = S_{\text{géométrie}} + S_{\text{matière}} \quad \text{avec} \quad S_{\text{géométrie}} = -\tfrac{c^3}{16\pi G}\int R\sqrt{\Delta}\,\mathrm{d}^4x\ .$$

Dans l'expression de $S_{\text{géométrie}}$, partie de l'action qui conduit aux équations d'Einstein dans le vide, $\sqrt{\Delta}\,\mathrm{d}^4x$ est un volume invariant (propriété qui se déduit de $|\det g'_{\alpha\beta}| = |\det g_{\mu\nu}|J^2$ où $J = \det(\partial x^\mu/\partial x'^\alpha)$ est le jacobien du changement de coordonnées, et $\Delta = |\det g_{\mu\nu}|$). $S_{\text{matière}}$ généralise l'action d'une particule ponctuelle $S_{\text{particule}} = \int -m\,c^2\,\mathrm{d}\tau = \int -m\,c\,\sqrt{g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu}$; enfin $T^{\mu\nu}$ est défini par

$$\delta S_{\text{matière}} = -\tfrac{1}{2c}\int T^{\mu\nu}\,\delta g_{\mu\nu}\,\sqrt{\Delta}\,\mathrm{d}^4x\ ,$$

relation semblable à $\delta(\int -mc^2\,\mathrm{d}\tau) = -\frac{1}{2c}\int mc^2 u^\mu u^\nu\,\delta g_{\mu\nu}\,c\,\mathrm{d}\tau$.

REMARQUE sur les dimensions. Comme en espace-temps plat $x^0 = ct$ et $\{x^i\} \equiv \vec{r}$, il est d'usage de poser $[x^\mu] = [L]$. On a alors $[g_{\mu\nu}] = [g^{\mu\nu}] = [\Delta] = 1$ (sans dimension), $[\Gamma^\beta_{\mu\nu}] = [L^{-1}]$, $[R^\beta_{\alpha\mu\nu}] = [R_{\alpha\mu}] = [R] = [L^{-2}]$ ... . Le lecteur vérifiera l'homogénéité des formules des sections 9.2.3 et 9.3.3.

Chapitre **10**

Probabilités ; processus aléatoires

Le calcul des probabilités intervient chaque fois que, répétant " à l'identique" une même expérience, le résultat trouvé pour la mesure d'une grandeur n'est pas unique : chiffre sur lequel un dé retombe, temps au bout duquel un noyau radioactif se désintègre, déplacement et vitesse d'une particule brownienne, mesure entachée d'erreur... Les raisons du caractère aléatoire du résultat sont multiples : sensibilité aux conditions initiales, interventions d'un grand nombre de variables dont la connaissance complète est impossible, mesures "classiques" sur des objets quantiques... Quelles que soient ces raisons le caractère aléatoire des résultats ne signifie pas l'absence d'information. L'approche probabiliste qui traite des fréquences d'apparition des résultats (idéalisées par les probabilités) permet le calcul de grandeurs moyennes et la confrontation à l'expérience. Elle joue un rôle fondamental dans la compréhension de nombreux phénomènes : loi des grands nombres et caractère gaussien des erreurs, distribution aléatoire d'instants et comptages obéissant à une loi de Poisson, marche au hasard et bruit blanc, équiprobabilité des états microscopiques et loi de Boltzmann en thermodynamique... Le calcul des probabilités joue également un rôle pratique important pour les problèmes d'estimation statistique.

10.1 LANGAGE DES PROBABILITÉS

Ce langage qui traite en apparence d'évidences (il y est question d'expériences et de leurs résultats) s'appuie sur des raisonnements de logique (utilisation de "ou", "et", "sachant que") qui demandent un minimum d'attention. Le concept clé est celui de probabilité d'un résultat. Appliqué aux grandeurs aléatoires il conduit à la notion de loi de probabilité (pour leurs valeurs), qui permet de définir les grandeurs moyennes.

10.1.1 Grandeurs aléatoires et raisonnements logiques ; conditionnement

▷ Expériences et grandeurs aléatoires

Le premier point à définir est l'ensemble Ω des expériences physiques ω considérées. Toutes doivent obéir à un même protocole (définition de l'état initial d'un échantillon radioactif, des règles d'un jeu de hasard...), et les mesures ou observations effectuées doivent être de même nature.

En général les expériences ω conduisent à des nombres (mesures de grandeurs physiques, gains attachés aux résultats d'un jeu, *etc.*). Une grandeur mesurée X est aléatoire si sa valeur $x = X(\omega)$ dépend de l'expérience $\omega \in \Omega$. On parle de **variable aléatoire (v.a.)** si $X(\omega)$ est un nombre, et de **vecteur aléatoire** (ou v.a. à n dimensions) si $\underline{X}$ est une famille $(X_1, X_2 \cdots X_n)$ de v.a. L'ensemble des valeurs possibles $X(\omega)$ (ou $\underline{X}(\omega)$) s'appelle le **spectre** de X (ou $\underline{X}$). Il peut être discret, continu ou même les deux à la fois (comme certains spectres d'énergie). Une v.a. qui dépend du temps est appelée **processus aléatoire**.

EXEMPLE (figure 1). Dans l'étude d'échantillons radioactifs identiques le nombre $N(t)$ de désintégrations dans l'intervalle $[0, t]$ est, à $t = \tau$ fixé, une v.a. discrète tandis que les instants T_1, T_2... des première, seconde... désintégrations sont des v.a. continues. A chaque expérience ω correspondent une suite d'instants $t_n = T_n(\omega)$ différente et un comptage $N(\tau, \omega)$ différent. On dit aussi que $N(t) = \sum_n H(t - T_n)$ est une **fonction aléatoire** dont les fonctions ordinaires $N(t, \omega) = \sum_n H(t - t_n)$ sont les réalisations (H fonction de Heaviside).

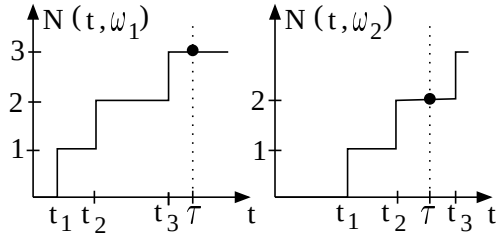


FIGURE 1

▷ Résultats et raisonnements logiques

Pour un protocole donné d'expérience on peut faire état de beaucoup de résultats (on dit aussi d'**évènements**). Ainsi, si il est vrai que pour deux lancers successifs d'un dé les résultats élémentaires directement observés sont les 36 couples ordonnés (i, j) avec $1 \leq i, j \leq 6$, on peut en imaginer d'autres, chacun étant défini par une affirmation $\{\cdots\}$ à laquelle une expérience répond nécessairement par "oui" ou par "non". Par exemple si X_1 et X_2 sont les nombres aléatoires du premier et du second lancer, $\{X_1 = 1\}$ est identique à $\{(1,1)\ ou\ (1,2) \ldots ou\ (1,6)\}$, $\{X_1 = i\ et\ X_2 = j\}$ est identique à $\{(i,j)\}$, $\{X_1 X_2 > 1\}$ est identique à $\{non\ (1,1)\}$, $\{X_1 X_2 = 7\}$ est impossible tandis que $\{X_1 X_2 \geq 1\}$ est certain.

Une représentation ensembliste (figure 2) des résultats (et opérations logiques) consiste à identifier chaque résultat A au sous ensemble de Ω (noté aussi A) dont les éléments sont les expériences conduisant ("répondant oui") à A. Alors aux résultats $\{A\ ou\ B\}$, $\{A\ et\ B\}$, $\{non\ A\}$ correspondent les sous ensembles $A \cup B$, $A \cap B$ et $\overline{A}$ (complémentaire de A). Les résultats certain et impossible correspondent à Ω $(= A \cup \overline{A})$ et $\emptyset$ $(= A \cap \overline{A})$. De très nombreux raisonnements font appel à la notion de **résultats exclusifs** tels que $\{A\ et\ B\}$ est impossible $(A \cap B = \emptyset)$, ainsi qu'à la notion de **partition**, famille de résultats exclusifs $(A_i \cap A_j = \emptyset$ pour $i \neq j)$ dont l'union est certaine $(\cup_i A_i = \Omega)$, par exemple A et $\overline{A}$. Une **partition** est **élémentaire** si tout résultat est une union de A_i. C'est le cas des résultats (i, j) ci-dessus, ou pour une v.a. discrète X prenant les valeurs x_i, des résultats $\{X = x_i\}$ (représentés par les sous ensembles $X^{-1}(x_i)$ de Ω ; figure 3). Pour une v.a. continue à valeurs dans $\mathbb{R}$, une partition de Ω est induite par une partition de $\mathbb{R}$ avec des intervalles pouvant "à la limite être infinitésimaux".

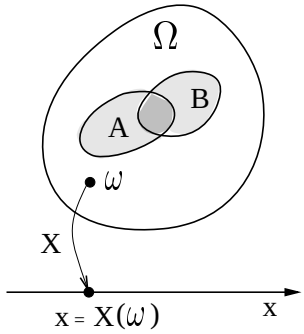


FIGURE 2

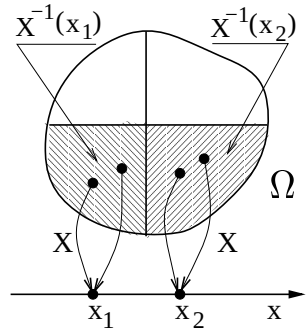


FIGURE 3

▷ Conditionnement

Cette notion logique très utilisée consiste simplement, si on sait qu'un résultat B est vérifié, à ne prendre en compte que les expériences qui conduisent à B et donc à remplacer Ω par $\Omega_B = B$. Le résultat $\{A\ sachant\ B\}$ noté A/B (A conditionné par B) correspond au sous ensemble $A \cap B$ de B. La connaissance d'un résultat peut changer radicalement les choses. Par exemple si le résultat B implique le résultat A, c.a.d $B \subset A$, le résultat A/B est certain tandis que B/A ne l'est pas, et cela bien que $A \cap B = B \cap A$ (exemple : $B \equiv \{X = 2\ ou\ 4\}$ et $A \equiv \{X\ pair\}$ pour le lancer d'un dé).

10.1.2 Probabilités ; lois de probabilité

▷ Probabilités de résultats

Quand on fait N expériences dont N_A ont donné le résultat A on observe que, pour N de plus en plus grand, le rapport $\dfrac{N_A}{N}$ tend vers une limite : la "fréquence d'apparition de A". L'approche probabiliste postule qu'à tout résultat A on peut associer un nombre, sa probabilité $P(A) \in [0, 1]$; il doit vérifier $P(\Omega) = 1$ (probabilité 1 pour un résultat certain) et $P(A\,ou\,B) = P(A) + P(B)$ si $\{A\,et\,B\} = \emptyset$: **addition des probabilités**

pour des résultats exclusifs (en particulier $\sum_i P(A_i) = 1$ pour une partition, d'où $P(\overline{A}) = 1 - P(A)$). Dans cette approche la donnée de P est une prédiction sur les fréquences des résultats ; ceci est nécessaire car on ne peut en effet pratiquement pas faire une infinité d'expériences.

▷ Probabilités conditionnelles ; indépendance

Si A est conditionné par B, la fréquence concernée est la limite du rapport $\dfrac{N_{A\,\text{et}\,B}}{N_B}$ qui est aussi le quotient de $\dfrac{N_{A\,\text{et}\,B}}{N}$ par $\dfrac{N_B}{N}$. C'est pourquoi on appelle probabilité conditionnelle la quantité :

$$P(A/B) = \frac{P(A\,et\,B)}{P(B)} \qquad \left(\neq P(B/A) = \frac{P(A\,et\,B)}{P(A)}\right) .$$

En particulier deux résultats sont dits indépendants si le conditionnement de l'un par l'autre est sans effet sur les probabilités :

$$P(A/B) = P(A) \text{ ou } P(B/A) = P(B) \Longleftrightarrow P(A\,et\,B) = P(A)\,P(B) .$$

Plus généralement on parle de résultats indépendants A_i si $P(A_1\,et\,A_2\,et\ldots) = P(A_1)\,P(A_2)\cdots$ (**multiplication des probabilités**). On notera qu'il ne faut pas confondre "exclusifs" (affirmation logique $\{A\,et\,B\} = \emptyset$) et "indépendants" (affirmation sur les probabilités).

EXEMPLE. Il est naturel de supposer que les résultats A_1 et A_2 associés à la non désintégration d'un noyau radioactif donné dans des intervalles $[t_1, t_1 + \Delta t_1]$ et $[t_2, t_2 + \Delta t_2]$ disjoints sont indépendants. Si on suppose de plus que $P(A_i)$ ne dépend que de Δt_i (et pas de t_i), c.a.d. si $P(A_i) = f(\Delta t_i)$ on obtient, en prenant $t_2 = t_1 + \Delta t_1$, $f(\Delta t_1 + \Delta t_2) = f(\Delta t_1)\,f(\Delta t_2)$ et donc $P(A) = e^{-\lambda\,\Delta t}$ (*cf.* section 1.1.3). Il est clair que A_1 et A_2 ne sont pas exclusifs. A l'opposé les désintégrations dans ces mêmes intervalles sont des résultats $\overline{A_1}$ et $\overline{A_2}$ exclusifs et non indépendants.

Les formules suivantes montrent comment, à partir des probabilités d'une partition B_i et des probabilités conditionnées par les B_i, on peut remonter à toutes les probabilités d'une autre partition A_j ainsi qu'aux probabilités conditionnelles "inverses" :

$$P(A) = \sum_i P(A/B_i)\,P(B_i) \quad ; \quad P(B_i/A_j) = \frac{P(A_j/B_i)\,P(B_i)}{\sum_i P(A_j/B_i)\,P(B_i)} .$$

(Il suffit d'appliquer les définitions et de remarquer que $A = \cup_i\,(A \cap B_i)$.) La première est très utilisée en physique. La seconde (**théorème de Bayes**) sert en analyse statistique.

▷ Lois de probabilité de v.a.

Si une v.a. X est discrète, à chaque valeur x_i est associée une probabilité $p_i \overset{\text{def}}{=} P\big(\{X = x_i\}\big)$ et on a $\sum_i p_i = 1$ puisque ces résultats forment une partition. Si la v.a. X est continue on appelle **densité de probabilité** la fonction $P_X(x)$ (dont la dimension est $[X]^{-1}$) telle que $P_X(x)\,\mathrm{d}x$ est la probabilité de $\{X \in [x, x + \mathrm{d}x]\}$; son intégrale sur $]-\infty, \infty[$ vaut 1.

EXEMPLE 1. La **loi de Bernouilli** correspond à une v.a. qui prend les valeurs 0 et 1 avec les probabilités $1-p$ et p. On peut aussi lui associer une densité $P_X(x) = (1-p)\,\delta(x) + p\,\delta(x-1)$ car chaque $p_i\,\delta(x-x_i)$ est localisé en x_i et a pour intégrale p_i (figure 4). Exemples : jeu de pile ou face, loi de Boltzmann pour un système à deux niveaux.

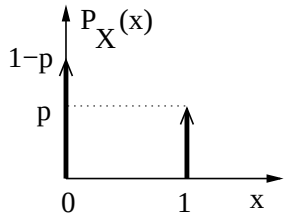


FIGURE 4

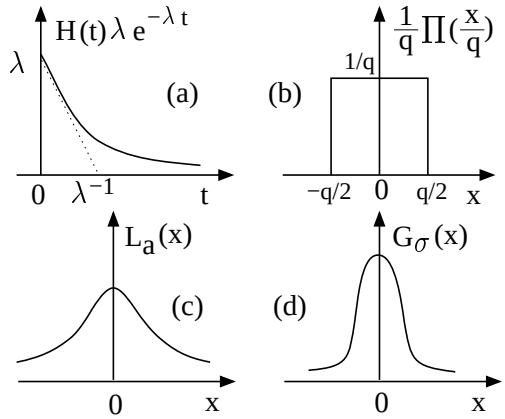


FIGURE 5

EXEMPLE 2. La densité de probabilité de l'instant T de désintégration d'un noyau est $P_T(t) = H(t)\,\lambda\,e^{-\lambda t}$ (**loi exponentielle** ; figure 5a). Démonstration : on écrit que la probabilité de $\{T < t\}$ est $\int_0^t P_T(t')\,\mathrm{d}t' = 1 - e^{-\lambda t}$ (car la probabilité de non désintégration dans l'intervalle $[0,t]$ est $e^{-\lambda t}$), et on dérive par rapport à t. Les densités d'autres lois continues, **loi uniforme**, **loi lorentzienne** (ou de Cauchy) $L_a(x)$ et **loi gaussienne** $G_\sigma(x)$, discutées plus loin, sont représentées sur les figures 5b,c,d.

La densité $P_Y(y)$ d'une fonction de v.a. $Y = g(X)$ s'obtient en remarquant que la probabilité $P_Y(y)\,\mathrm{d}y$ de $\{Y \in [y, y+\mathrm{d}y]\}$ est celle de $\{X \in [x_1, x_1+\mathrm{d}x_1]$ *ou* $[x_2, x_2+\mathrm{d}x_2]$ *ou* $\cdots\}$ où les x_i constituent l'image inverse $g^{-1}(y)$. Par exemple (figure 6) pour $Y = a\cos X$ et $P_X(x) = \dfrac{1}{2\pi}$ sur $[-\pi,\pi]$ on a $|\mathrm{d}x_1| = |\mathrm{d}x_2| = |\mathrm{d}y|\,(a^2-y^2)^{-\frac{1}{2}}$ et $P_Y(y) = \displaystyle\sum_i P_X(x_i)\left|\frac{\mathrm{d}x_i}{\mathrm{d}y}\right| = \frac{1}{\pi}\,(a^2-y^2)^{-\frac{1}{2}}$.

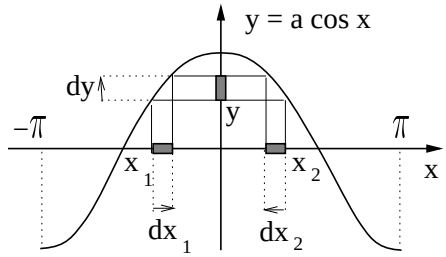


FIGURE 6

▷ Lois de probabilité relatives à une famille de v.a.

On appelle **densité de probabilité conjointe** de n v.a. la fonction $P_{\underline{X}}(\underline{x})$ telle que

$$P_{\underline{X}}(\underline{x})\,\mathrm{d}x_1 \cdots \mathrm{d}x_n$$

est la probabilité de $\{X_1 \in [x_1, x_1 + \mathrm{d}x_1]\ et \cdots et\ X_n \in [x_n, x_n + \mathrm{d}x_n]\}$. Si on s'intéresse à une seule variable, X_1 par exemple, sa **densité de probabilité marginale** $P_{X_1}(x_1)$ s'obtient (les autres v.a. pouvant prendre n'importe quelle valeur) en intégrant $P_{\underline{X}}(\underline{x})$ sur $x_2 \cdots x_n$:

$$P_{X_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P_{\underline{X}}(x_1, x_2, \cdots, x_n)\,\mathrm{d}x_2 \cdots \mathrm{d}x_n \ ;$$

sa **densité de probabilité conditionnelle** (conditionnée par les valeurs $x_2 \cdots x_n$ des autres v.a.) s'obtient en faisant le quotient de la probabilité conjointe $P_{\underline{X}}(\underline{x})\,\mathrm{d}x_1\,\mathrm{d}x_2 \cdots \mathrm{d}x_n$ par la probabilité de la condition $\mathrm{d}x_2 \cdots \mathrm{d}x_n \int P_{\underline{X}}(\underline{x})\,\mathrm{d}x_1$, c.a.d. simplement en normalisant la fonction de x_1 obtenue en fixant $x_2 \cdots x_n$ dans $P_{\underline{X}}(\underline{x})$:

$$P_{X_1}(x_1/x_2 \cdots x_n) = \frac{P_{\underline{X}}(x_1, x_2 \cdots x_n)}{\int P_{\underline{X}}(x_1, x_2 \cdots x_n)\,\mathrm{d}x_1} \qquad (x_2 \cdots x_n \text{ fixés}) \ .$$

Pour des v.a. indépendantes la densité de probabilité conjointe est le produit des densités marginales : $P_{\underline{X}}(\underline{x}) = P_{X_1}(x_1) \cdots P_{X_n}(x_n)$. Enfin la densité de probabilité $P_Y(y)$ d'une fonction de plusieurs v.a. $Y = g(\underline{X})$ s'obtient en intégrant $P_{\underline{X}}(\underline{x})$ sur le domaine des valeurs de $\underline{X}$ correspondant à $\{Y \in [y, y + \mathrm{d}y]\}$ et en divisant le résultat par $\mathrm{d}y$.

EXEMPLE 1. La figure 7a correspond au cas de deux v.a. X et Y avec une densité conjointe uniforme sur le disque $x^2 + y^2 \leq 1$. La figure 7b montre la densité marginale $P_X(x) = \dfrac{2}{\pi}\sqrt{1 - x^2}$ de X (identique à celle de Y). Comme $P_{\underline{X}}(x, y) \neq P_X(x)\,P_Y(y)$, ces v.a. ne sont pas indépendantes (elles le seraient dans le cas d'un domaine rectangulaire de côtés parallèles aux axes x et y). Ceci est confirmé par le fait que les densités conditionnelles $P_X(x/y) = \dfrac{1}{2}\,(1 - y^2)^{-\frac{1}{2}}$ (pour $|x| < (1 - y^2)^{\frac{1}{2}}$) de X dépendent de y (figure 7c).

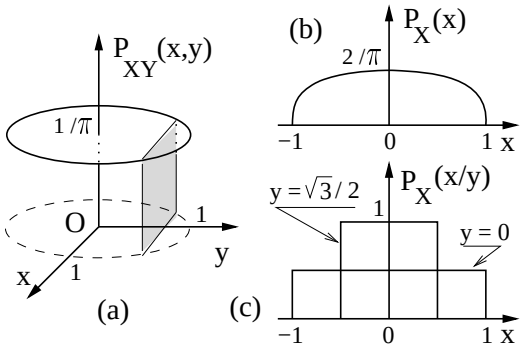


FIGURE 7

EXEMPLE 2 (figure 8). La densité de probabilité du vecteur vitesse $\overrightarrow{V} \equiv (V_x, V_y, V_z)$ d'une molécule dans un fluide à l'équilibre à une température fixée ($\beta = (k_B T)^{-1}$) est

$$P_{\overrightarrow{V}}(\vec{v}) = Z^{-1} \exp\left(-\beta \frac{1}{2} m\vec{v}^2\right)$$

(**distribution de Maxwell**), où $Z = \left(\frac{2\pi}{m\beta}\right)^{\frac{3}{2}}$ est donné par $\iiint P_{\overrightarrow{V}}(\vec{v})\,\mathrm{d}^3 v = 1$. Les trois composantes de $\overrightarrow{V}$ sont donc indépendantes et la densité (marginale) de V_x par exemple est :

$$P_{V_x}(v_x) = \left(\frac{2\pi}{m\beta}\right)^{-\frac{1}{2}} \exp\left(-\beta \frac{1}{2} m v_x^2\right) .$$

Les densités de $V = |\overrightarrow{V}| > 0$ et $E = \frac{1}{2} mV^2 > 0$, fonctions de $\overrightarrow{V}$, sont respectivement

$$P_V(v) = Z^{-1} 4\pi v^2 \exp\left(-\beta \frac{1}{2} m\vec{v}^2\right)$$

(obtenue à partir de l'intégration de $P_{\overrightarrow{V}}(\vec{v})$ sur le domaine $|\vec{v}| \in [v, v + \mathrm{d}v]$) et

$$P_E(\epsilon) = 2\pi^{-\frac{1}{2}} \beta^{\frac{3}{2}} \sqrt{\epsilon} \exp -\beta\epsilon$$

(obtenue en écrivant $P_E(\epsilon)\,\mathrm{d}\epsilon = P_V(v)\,\mathrm{d}v$ avec $\epsilon = \frac{1}{2} mv^2$ soit $\mathrm{d}v = (2m\epsilon)^{-\frac{1}{2}}\,\mathrm{d}\epsilon$).

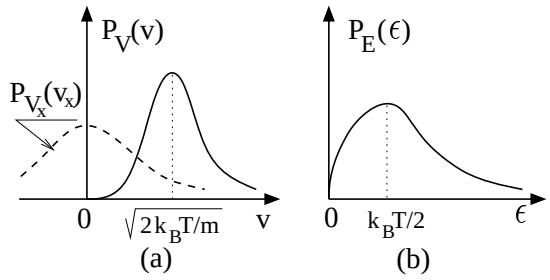


FIGURE 8

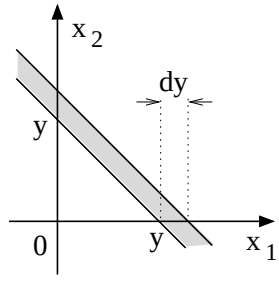


FIGURE 9

EXEMPLE 3. Pour une **lumière thermique** la densité conjointe des parties réelle X et imaginaire Y de l'amplitude complexe $A = X + iY$ est $P_{X,Y}(x, y) = (2\pi\sigma^2)^{-1} \exp\left(-\frac{1}{2\sigma^2}(x^2 + y^2)\right)$. On en déduit celle de l'intensité $\mathcal{I}$ et de la phase Φ ($A = \sqrt{\mathcal{I}} e^{i\Phi}$) en écrivant

$$P_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y = P_{\mathcal{I},\Phi}(I, \varphi)\,\mathrm{d}I\,\mathrm{d}\varphi \quad \text{avec} \quad \mathrm{d}x\,\mathrm{d}y = \frac{1}{2}\,\mathrm{d}I\,\mathrm{d}\varphi\ ;$$

on obtient $P_{\mathcal{I},\Phi}(I, \varphi) = \frac{1}{2\pi} \frac{1}{2\sigma^2} \exp\left(-\frac{I}{2\sigma^2}\right)$. La densité de $\mathcal{I}$ est $P_{\mathcal{I}}(I) = \frac{1}{2\sigma^2} \times \exp\left(-\frac{I}{2\sigma^2}\right)$ (loi exponentielle), celle de Φ est $P_\Phi(\varphi) = \frac{1}{2\pi}$ (loi uniforme sur $[0, 2\pi]$) et les deux v.a. sont indépendantes.

EXEMPLE 4. **Somme de deux v.a.** $Y = X_1 + X_2$. L'observation de la figure 9 montre que

$$P_Y(y) = \int_{-\infty}^{\infty} P_{X_1,X_2}(x_1, y - x_1)\,\mathrm{d}x_1 \ .$$

Si les v.a. sont indépendantes $P_{X_1,X_2}(x_1, y - x_1) = P_{X_1}(x_1)\,P_{X_2}(y - x_1)$; donc la densité de probabilité $P_{X_1+X_2}$ de la somme de deux v.a. indépendantes est le **produit de convolution** $P_{X_1} * P_{X_2}$ des deux densités.

10.1.3 Grandeurs moyennes ; moments ; corrélations

▷ Cas d'une v.a.

Si pour une v.a. X discrète une valeur x_i est obtenue N_i fois lors de N expériences, la moyenne "expérimentale " est $N^{-1}\sum_i x_i\,N_i$; la moyenne "théorique" est $< X >= \sum_i x_i\,p_i$. (La notation $< \cdots >$ pour la moyenne est celle couramment utilisée en physique ; en mathématique on écrit plutôt $E(\cdots)$ et le mot moyenne est remplacé par "**espérance**", "expectation" en anglais).

De façon générale on appelle **moyenne** d'une fonction $f(X)$ de la v.a. X continue (ou discrète) le nombre

$$< f(X) >= \int_{-\infty}^{\infty} f(x)\,P_X(x)\,\mathrm{d}x \quad (\text{ou } \sum_i f(x_i)\,p_i) \ .$$

En particulier la moyenne de X est

$$m_X =< X >= \int_{-\infty}^{\infty} x\,P_X(x)\,\mathrm{d}x \quad (\text{ou } \sum_i x_i\,p_i \ )$$

($m_X = 0$ pour une **v.a. centrée**), et sa **variance** est le nombre positif

$$\sigma_X^2 =< (X - m_X)^2 >= \int_{-\infty}^{\infty} (x - m_X)^2\,P_X(x)\,\mathrm{d}x =< X^2 > -m_X^2$$

($\sigma_X^2 = 1$ pour une **v.a. réduite**). On introduit aussi la **fonction caractéristique** (reliée à la T.F. de $P_X(x)$) :

$$< e^{iuX} >= \int_{-\infty}^{\infty} e^{iux}\,P_X(x)\,\mathrm{d}x \quad (\text{ou } \sum_i e^{iux_i}\,p_i) \ .$$

Lorsque cette fonction de u admet un développement de Taylor en $u = 0$, on obtient tous les **moments d'ordre n** par

$$< e^{iuX} >= \sum_n \frac{(iu)^n}{n!} \ < X^n >$$

(idem pour les **moments centrés** $< (X - m_X)^n >$ à partir du développement de $< e^{iu(X - m_X)} >$).

EXEMPLES (figures 4 et 5b,c,d). **Loi de Bernouilli** (*cf.* section 10.1.2) : $< X >= p$, $\sigma_X^2 = p(1-p)$ et $< e^{iuX} >= pe^{iu} + q$ (avec $q = 1 - p$).

Loi uniforme sur $\left[-\frac{q}{2}, \frac{q}{2}\right]$: $P_X(x) = \frac{1}{q}\,\Pi\Big(\frac{x}{q}\Big)$, $< X >= 0$ et $\sigma_X^2 = \frac{q^2}{12}$.

Loi gaussienne :

$P_X(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp -\dfrac{(x-m)^2}{2\sigma^2} \Longleftrightarrow < e^{iuX} >= e^{ium} e^{-\frac{u^2\sigma^2}{2}}$; $< X >= m$ et $\sigma_X^2 = \sigma^2$.

Le développement de $\exp -\dfrac{u^2\sigma^2}{2}$ montre que les seuls moments centrés non nuls sont $< (X-m)^{2p} >= 1.3.5 \cdots (2p-1)\, \sigma^{2p}$; en particulier $< (X-m)^4 >= 3 < (X-m)^2 >^2$.

REMARQUES. La variance, lorsqu'elle existe, " mesure" la tendance d'une v.a. à être plus ou moins "piquée" autour de sa moyenne. En effet on démontre que la probabilité pour que X s'écarte de sa moyenne de plus de λ fois l'**écart type** σ_X est inférieure à λ^{-2} ; pour une v.a. gaussienne elle vaut 5×10^{-2} pour $\lambda = 2$, et 4×10^{-3} pour $\lambda = 3$. Si la variance est nulle la v.a. est certaine (et réciproquement). Il existe cependant des exemples comme la loi lorentzienne $P_X(x) = \dfrac{a}{\pi} \dfrac{1}{a^2 + x^2}$, présentant une "largeur typique a", mais n'ayant pas d'écart type ; ($< e^{iuX} >= e^{-a|u|}$ n'est pas non plus développable en puissances de u).

▷ Cas d'une famille de v.a.

La définition des moyennes est une généralisation de celle donnée pour une v.a. :

$$< f(\underline{X}) >= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\underline{x})\, P_X(\underline{x})\, \mathrm{d}x_1 \cdots \mathrm{d}x_n \ .$$

Donc la moyenne d'une somme de fonctions des v.a. est la somme des moyennes de chaque fonction. En plus des moyennes $m_i =< X_i >$ de chaque v.a. et de leurs variances $\sigma_i^2 =< (X_i - m_i)^2 >$, les quantités les plus utilisées sont les **corrélations** :

$$\Gamma_{ij} =< (X_i - m_i)(X_j - m_j) >=< X_i X_j > - m_i m_j \quad (i \neq j \ ; \ \Gamma_{ii} = \sigma_i^2) \ .$$

Γ_{ij} mesure la tendance des v.a. centrées $X_i - m_i$ et $X_j - m_j$ à varier dans le même sens ($\Gamma_{ij} > 0$) ou en sens inverse ($\Gamma_{ij} < 0$). Les v.a. X_i et X_j sont corrélées (resp. non corrélées) si $\Gamma_{ij} \neq 0$ (resp. $\Gamma_{ij} = 0$).

Dans le cas de v.a. indépendantes $P_{\underline{X}}(\underline{x})$ factorise. Donc si $f(\underline{X})$ factorise aussi on a :

$$< \Pi_i f_i(X_i) >= \Pi_i < f_i(X_i) > \ .$$

En particulier $< X_i X_j >=< X_i >< X_j >$ d'où $\Gamma_{ij} = 0$ $(i \neq j)$. Donc l'indépendance des X_i entraîne leur non corrélation. On fera attention que, sauf pour les v.a. gaussiennes, la réciproque est fausse. (*cf.* exemple 1 de la section 10.1.2. pour lequel il est évident que $< X >=< Y >=< XY >= 0$ alors que X et Y ne sont pas indépendantes).

10.2 ORIGINE ET DISCUSSION DE QUELQUES LOIS IMPORTANTES EN PHYSIQUE

10.2.1 Théorème de la limite centrale et lois gaussiennes

Dans de nombreuses situations physiques la v.a. X est la somme d'un grand nombre de v.a. X_i. Ce peut être une erreur X accumulation de petites erreurs, le déplacement $\overrightarrow{R} = \overrightarrow{R}_1 + \overrightarrow{R}_2 + \cdots + \overrightarrow{R}_N$ d'une marche au hasard constituée de N pas, la somme

$\sum_i |A_i|\, e^{i\Phi_i}$ de nombres complexes (amplitude complexe d'une lumière thermique ou d'une figure de speckles), un signal résultant $X(t) = \sum_i f(t - T_i)$ (à t fixé) lorsque beaucoup de fonctions $f(t - t_i)$ se recouvrent (superposition de paquets d'ondes émis à des instants aléatoires), *etc.* On observe alors généralement que la loi de X est bien approchée par une gaussienne. On va le justifier dans le cas où les X_i sont des v.a. indépendantes et de même loi, caractérisée par une moyenne m et une variance σ^2. Ces conditions peuvent être "assouplies" ; l'existence des variances et de "pas trop de corrélation" suffit.

▷ Propriétés de $\mathbf{X = X_1 + X_2 + \cdots + X_N}$

La linéarité de l'opération de moyenne entraîne

$$< X >= \sum_{i=1}^{N} < X_i > \quad ; \quad < X^2 >= \sum_{i,j=1}^{N} < X_i X_j > \ .$$

Si les X_i sont indépendantes on a $< X_i X_j >=< X_i >< X_j >$ pour $i \neq j$. On en déduit que la variance de X est la somme des variances des X_i et que si les X_i ont même loi :

$$< X >= Nm \quad \text{et} \quad \sigma_X^2 = N\sigma^2 \ .$$

EXEMPLES. Si les X_i sont des **erreurs** de moyenne nulle et de valeur "typique" σ, l'erreur X à une valeur "typique" $\sqrt{N}\sigma$. Dans une **marche au hasard** sans dérive ($< \overrightarrow{R}_i >= 0$), le "déplacement typique" au bout de N pas est $\sqrt{N}l$ avec $l = \sqrt{< \overrightarrow{R}_i^{\,2} >}$. Enfin la v.a. $\dfrac{X}{N}$ qui a pour écart type $\dfrac{\sigma}{\sqrt{N}}$ devient la variable sure égale à m pour $N \to \infty$.

REMARQUE. Si $m \neq 0$ on peut assimiler X à sa valeur moyenne Nm car l'erreur relative commise $\dfrac{\sigma_X}{< X >} = \dfrac{\sigma}{m\sqrt{N}}$ tend vers 0 pour N grand ; c'est le cas des **grandeurs extensives** de la thermodynamique. Par contre ceci devient faux pour des sommes de v.a. "fortement dépendantes", par exemple $\sigma_X^2 = N^2\sigma^2$ pour $X = X_1 + \cdots + X_1 = NX_1$. On peut aussi vérifier explicitement que pour l'intensité d'une **figure de speckles** $\mathcal{I} = \left|\sum_i e^{i\Phi_i}\right|^2 = \sum_{i,j=1}^{N} e^{i(\Phi_i - \Phi_j)}$, avec Φ_i v.a. indépendantes et uniformes sur $[0, 2\pi]$, on a $< \mathcal{I} >= N$ et (petit calcul) $\sigma_{\mathcal{I}}^2 = N^2 - N$; la figure présente de grandes fluctuations d'intensité.

▷ Théorème de la limite centrale

Lorsque N tend vers l'infini la v.a. $Z = \dfrac{X - Nm}{\sqrt{N}}$, qui est centrée et de variance σ^2, tend vers une gaussienne. On peut alors approximer la densité de la v.a. $X = \sum_{i=1}^{N} X_i$, pour N grand, par

$$P_X(x) \simeq \frac{1}{\sqrt{2\pi N\sigma^2}}\ \exp -\frac{(x - Nm)^2}{2N\sigma^2} \ .$$

JUSTIFICATION : Z est la somme des N v.a. $X_i' = \frac{1}{\sqrt{N}}(X_i - m)$ qui sont centrées et de variance $\frac{\sigma^2}{N}$. Sa fonction caractéristique $< e^{iuZ} >$ est le produit $\Pi_i < e^{iuX_i'} >$ (indépendance des X_i'), égal à $< e^{iuX_1'} >^N$ (X_i' de même loi), soit $\left(1 - \frac{\sigma^2}{2N}\, u^2 + \cdots\right)^N$ qui tend vers $\exp -\frac{\sigma^2 u^2}{2}$ (si la somme des termes non écrits du développement peut être négligée).

REMARQUES. 1) Si les X_i sont équiréparties sur $\left[-\frac{1}{2},\frac{1}{2}\right]$ leur densité est $\Pi(x)$, celle de $N^{-\frac{1}{2}}X_i$ est $\sqrt{N}\,\Pi(\sqrt{N}x)$ et celle de Z est le produit de convolution $\left(\sqrt{N}\Pi(\sqrt{N}x)\right)^{*N}$; pour $N\to\infty$ il tend vers une gaussienne comme annoncé à la section 5.2.1.
2) Si les X_i sont des v.a. lorentziennes de densité $L_a(x)$ la densité de X est $L_{Na}(x)$ (évident en considérant les fonctions caractéristiques) qui ne tend pas vers une gaussienne pour N grand ; ceci ne contredit pas le théorème car alors σ^2 n'existe pas.

▷ Famille de v.a. gaussiennes

On a déjà vu, avec les composantes de la vitesse d'une molécule d'un gaz, un exemple de trois v.a. gaussiennes indépendantes. Dans le cas général une loi gaussienne (en supposant les variables centrées pour simplifier l'écriture) est définie par la densité (**A** matrice symétrique positive)

$$P_{\underline{X}}(\underline{x}) = (2\pi)^{-\frac{n}{2}}\,(\det\mathbf{A})^{\frac{1}{2}}\,\exp -\frac{1}{2}\underline{x}^t\mathbf{A}\underline{x} \qquad (\underline{x}^t\mathbf{A}\underline{x} = \sum_{i,j=1}^{n} x_i A_{ij} x_j)\ ,$$

ou de façon équivalente par la fonction caractéristique :

$$< e^{i\underline{u}^t\underline{X}} >= \exp -\frac{1}{2}\,\underline{u}^t\mathbf{\Gamma}\underline{u} \qquad \text{avec } \mathbf{A}=\mathbf{\Gamma}^{-1} \qquad (\underline{u}^t\underline{X} = \sum_{i=1}^{n} u_i X_i)\ .$$

L'équivalence des deux définitions est facile à démontrer dans la base où **A** est diagonale.

La comparaison des développements au second ordre en $\underline{u}$ de $< e^{i\underline{u}^t\underline{X}} >$ et $e^{-\frac{1}{2}\underline{u}^t\mathbf{\Gamma}\underline{u}}$ montre que $\Gamma_{ij} =< X_iX_j >$. Si les X_i ne sont pas corrélées, $\mathbf{\Gamma}$ et donc aussi **A** sont diagonales et la densité factorise ; on en déduit que **non corrélation** et **indépendance** sont des notions identiques pour des v.a. gaussiennes. Une autre propriété importante est que toute combinaison linéaire de v.a. gaussiennes $Y=\sum_i a_iX_i$ est aussi gaussienne ; en effet $< e^{iuY} >= \exp -\frac{1}{2}u^2\sum_{i,j} a_i\Gamma_{ij}a_j$ (car $u_i = ua_i$) est la fonction caractéristique d'une gaussienne de variance $\sigma_Y^2 = \sum_{i,j} a_i\Gamma_{ij}a_j$. Plus généralement le filtrage linéaire préserve le caractère gaussien d'un processus.

REMARQUE. Pour des v.a. non centrées il suffit de remplacer dans ce qui précède $\underline{X}$ par $\underline{X}-\underline{m}$ avec $m_i =< X_i >$.

10.2.2 Loi binomiale et loi de Poisson

▷ Loi binomiale

C'est la loi de probabilité de la somme $X=X_1+X_2+\cdots+X_N$ de N v.a. de Bernouilli indépendantes et de même loi (caractérisée par la probabilité p). Donc :

$$< X >= N < X_1 >= Np \ \text{ et } \ \sigma_X^2 = N\sigma_{X_1}^2 = Np(1-p)\ .$$

De la fonction caractéristique $< e^{iuX} >= (pe^{iu}+q)^N = \sum_{n=0}^{N} C_N^n\,p^n\,q^{N-n}\,e^{iun} = \sum_{n=0}^{N} p_n\,e^{iun}$, on déduit que X prend les valeurs $n=0,1\cdots N$ avec les probabilités (caractérisées par N et p) :

$$p_n = C_N^n\,p^n\,q^{N-n} \qquad (q=1-p)\ .$$

On peut aussi obtenir ce résultat en remarquant qu'il y a C_N^n façons d'avoir n v.a. égales à 1 et $N-n$ v.a. égales à 0 et que pour chacune la probabilité est le produit $p^n q^{N-n}$.

EXEMPLE. Dans une enceinte de volume V contenant N molécules d'un gaz parfait, X compte le nombre de molécules dans un volume ΔV : $X_i = 1$ ou 0 selon la présence ou l'absence de la molécule i dans ΔV et $p = \dfrac{\Delta V}{V}$. La fluctuation relative de densité $\dfrac{\sigma_X}{< X >} = \Big(\dfrac{V-\Delta V}{N\Delta V}\Big)^{\frac{1}{2}}$ est négligeable si ΔV reste macroscopique ; on peut alors dire que, à l'échelle ΔV, la répartition des particules dans V est uniforme.

▷ Loi de Poisson

Elle est la limite de la loi binomiale lorsque N tend vers l'infini, p tend vers zéro, le produit $Np = m$ restant fini. Le spectre d'une v.a. de Poisson est donc constitué par tous les entiers $n = 0, 1, 2 \cdots$ et :

$$< X >= \sigma_X^2 = m \quad ; \quad p_n = \frac{m^n\, e^{-m}}{n!} \quad ; \quad < e^{iuX} >= e^{m(e^{iu}-1)} \ .$$

DÉMONSTRATION : on récrit pour la loi binomiale

$$p_n = \frac{N(N-1)\dots(N-n+1)}{n!}\Big(\frac{m}{N}\Big)^n\Big(1-\frac{m}{N}\Big)^N\Big(1-\frac{m}{N}\Big)^{-n} \ ,$$

$$< e^{iuX} >= \Big(1+(e^{iu}-1)\,\frac{m}{N}\Big)^N \ ,$$

puis on fait tendre N vers l'infini. La loi de Poisson ne dépend que du seul paramètre m.

REMARQUE. En faisant le produit des fonctions caractéristiques, on voit que la somme de v.a. de Poisson indépendantes de moyenne m_i est une v.a. de Poisson de moyenne $\sum_i m_i$. Une v.a. de Poisson dont la moyenne est grande ($m \gg 1$) est quasi-continue ; elle peut être vue comme la somme d'un grand nombre de v.a de Poisson et (théorème de la limite centrale) elle se comporte donc comme une v.a. gaussienne ; son écart type n'est pas quelconque mais relié à la moyenne par $\sigma = \sqrt{m}$.

▷ Evènements aléatoires dans le temps

La **loi de Poisson** intervient chaque fois que des évènements se produisant de façon aléatoire dans le temps obéissent aux trois hypothèses suivantes (très souvent satisfaites expérimentalement) : *i*) l'occurrence de plus d'un évènement dans un intervalle dt a une probabilité négligeable (pas d'évènements simultanés) ; *ii*) la probabilité d'occurrence d'un évènement dans un intervalle dt est $\lambda\,$dt (λ constant indépendant de t) ; *iii*) les occurrences d'évènements qui se produisent dans des intervalles de temps qui ne se recouvrent pas sont indépendantes. En effet en divisant un intervalle de largeur Δt en un grand nombre N d'intervalles de largeur $\delta t = \dfrac{\Delta t}{N}$, on obtient pour la probabilité $P_n(\Delta t)$ de l'occurence de n évènements dans Δt (limite $N \to \infty$ de $C_N^n\big(\dfrac{\lambda\Delta t}{N}\big)^n\big(1-\dfrac{\lambda\Delta t}{N}\big)^{N-n}$) :

$$P_n(\Delta t) = \frac{(\lambda\Delta t)^n}{n!}\, e^{-\lambda\Delta t} \ .$$

REMARQUES. 1) De l'expression de $P_n(\Delta t)$ on déduit que la probabilité pour que le n$^{\text{ième}}$ évènement, compté à partir de $t = 0$, se produise dans l'intervalle $[t, t + \mathrm{d}t]$ est $P_{n-1}(t)\, P_1(\mathrm{d}t) = \frac{(\lambda t)^{n-1}}{(n-1)!}\, e^{-\lambda t}\, \lambda\, \mathrm{d}t$. Donc la densité de probabilité de la v.a. T_n introduite à la section 10.1.1 est :

$$P_{T_n}(t_n) = \tfrac{(\lambda t_n)^{n-1}}{(n-1)!}\, e^{-\lambda t_n}\, \lambda\ .$$

Pour $n = 1$ on retrouve la loi exponentielle, et comme l'origine des temps est arbitraire, cette loi est aussi valable pour l'intervalle de temps $T_i - T_{i-1}$ entre deux désintégrations. De l'indépendance des v.a. $T_1, T_2 - T_1 \cdots T_n - T_{n-1}$, on déduit que la densité de probabilité conjointe des instants $T_1, T_2 \cdots T_n$ des n premiers évènements est le produit :

$$P_{T_1,T_2\cdots T_n}(t_1, t_2 \cdots t_n) = \lambda H(t_1) e^{-\lambda t_1}\, \lambda H(t_2 - t_1) e^{-\lambda(t_2-t_1)} \ldots \lambda H(t_n - t_{n-1}) e^{-\lambda(t_n - t_{n-1})}\ .$$

2) Si λ dépend du temps, la remarque sur l'addition des v.a. de Poisson montre qu'il suffit de remplacer $\lambda\Delta t$ par $\int_t^{t+\Delta t} \lambda(t')\, \mathrm{d}t'$ dans les expressions précédentes (réflexion laissée au lecteur).

Ces résultats s'appliquent par exemple aux émissions de particules par un échantillon radioactif, aux collisions d'un atome dans un gaz, aux appels qui arrivent à un standard téléphonique, *etc.* ainsi qu'aux signaux aléatoires du type $\sum_i f(t - T_i)$ (bruit de grenaille ; figure 10a), ou plus généralement aux processus qui consistent, aux instants aléatoires T_i, à effectuer une modification (souvent elle aussi aléatoire) du signal.

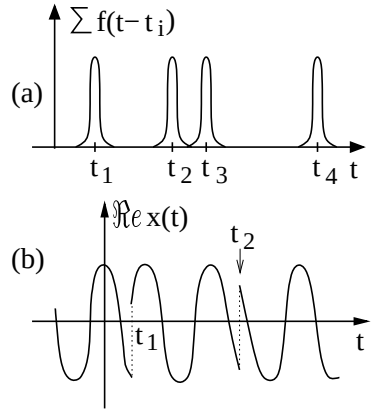


FIGURE 10

EXEMPLE : **élargissement de raie**. Le signal lumineux $a\, \exp -i(2\pi\nu_0 t + \varphi_0)\, \exp -\gamma t$ émis à $t = 0$ par un atome non perturbé (γ étant lié à la largeur naturelle de la raie) voit sa phase modifiée d'une quantité aléatoire à chaque collision avec un autre atome (fig.10b avec $\gamma = 0$). A un instant $t > 0$ le signal est $X(t) = a\, \exp -i\big(2\pi\nu_0 t + \varphi_0 + \Phi(t)\big)\, \exp -\gamma t$, avec $\Phi(t) = 0$ si il n'y a pas eu de choc entre 0 et t (probabilité $e^{-\lambda t}$), et avec $\Phi(t)$ équirépartie sur $[0, 2\pi]$ si il y a eu un ou plusieurs chocs (probabilité $1 - e^{-\lambda t}$). On voit sur le "signal moyen" $e^{-\lambda t}\, (a\, \exp -i(2\pi\nu_0 t + \varphi_0)\, \exp -\gamma t) + (1 - e^{-\lambda t}) \times 0$ que l'effet des collisions est d'élargir la raie en changeant γ en $\gamma + \lambda$ (*cf.* aussi section 10.3.3).

10.2.3 Loi de Boltzmann ; généralisations ; statistiques quantiques ; réponse linéaire et fluctuations

On se limite à l'établissement et à quelques propriétés générales de la loi de Boltzmann (*cf.* aussi section 7.4.2). Le lecteur est renvoyé aux ouvrages de thermodynamique et de physique statistique pour les applications.

▷ Etablissement de la loi de Boltzmann

On considère pour commencer un système macroscopique isolé $\mathcal{S}$ d'énergie U_0 fixée constitué de N_0 particules discernables sans interaction, chacune d'elle étant susceptible d'être dans des états α (discrets pour simplifier l'écriture) bien définis d'énergie ϵ_α. On suppose que les états microscopiques possibles de $\mathcal{S}$, au nombre de $\exp\left(k_B^{-1}\, S(U_0, N_0)\right)$ d'après la définition de l'entropie (*cf.* section 1.1.4), sont équiprobables (*cf.* section 10.3.2 "distribution microcanonique"). La probabilité p_α pour qu'une particule donnée soit dans l'état α (et donc les N_0-1 autres dans des états quelconques mais avec une énergie totale $U_0-\epsilon_\alpha$) s'obtient alors en faisant le rapport du "nombre de cas favorables" au "nombre de cas possibles". En utilisant la formule générale de thermodynamique $\mathrm{d}S = T(\mathrm{d}U - \mu\,\mathrm{d}N)$ ($\mu = \dfrac{F}{N}$, avec $F = U - TS$, est le potentiel chimique par particule) et l'inégalité $\epsilon_\alpha \ll U_0$, un développement au premier ordre donne :

$$p_\alpha = \frac{\exp\left(k_B^{-1}\, S(U_0 - \epsilon_\alpha, N_0 - 1)\right)}{\exp\left(k_B^{-1}\, S(U_0, N_0)\right)} = \exp -\frac{\epsilon_\alpha - \mu}{k_B T}\ .$$

Une autre écriture, pratique pour les calculs, utilise la **fonction de partition** $Z(\beta)$:

$$p_\alpha = \frac{e^{-\beta\epsilon_\alpha}}{Z(\beta)} \quad \text{avec} \quad Z(\beta) = \sum_\alpha e^{-\beta\epsilon_\alpha} = e^{-\beta\mu} \quad \left(\beta = \frac{1}{k_B T}\right).$$

REMARQUE. Si g_i est le nombre d'états d'énergie ϵ_i (g_i **facteur de dégénérescence** du niveau ϵ_i) la probabilité qu'une particule ait l'énergie ϵ_i est $p(\epsilon_i) = p_i = Z^{-1}\, g_i\, e^{-\beta\epsilon_i}$.

Particules en interaction (interactions supposées à courte portée). Au lieu de considérer les particules individuellement, on divise par la pensée le système $\mathcal{S}$ (de volume V_0) en un grand nombre de sous systèmes de volume V ($V \ll V_0$) comportant chacun un nombre suffisant N de particules ($N \ll N_0$ mais macroscopique) pour que les interactions entre sous systèmes (effet de surface) soient négligeables devant les interactions internes (effet de volume). Alors la démonstration précédente, reprise pour les sous systèmes et leurs états A, conduit à la loi de Boltzmann $p_A \propto e^{-\beta E_A}$ si on ignore les fluctuations de N. Par exemple dans un gaz réel la densité de probabilité conjointe des vitesses et positions des particules est $P(\vec{v}_1, \cdots \vec{v}_N; \vec{r}_1, \cdots \vec{r}_N) \propto \exp -\beta\Big(\sum_i \frac{1}{2}\, m\vec{v}_i^{\,2} + \sum_{(i,j)}\ V(\vec{r}_i - \vec{r}_j)\Big)$; contrairement au cas du gaz parfait, les positions ne sont pas indépendantes ; par contre les vitesses le restent.

REMARQUE. Les niveaux d'énergie des sous systèmes ont une **dégénérescence "macroscopique"** :

$$g(E) = \exp\tfrac{S(E,N)}{k_B} \ \text{ et donc } \ p(E) \propto \exp\tfrac{S(E,N)}{k_B}\, e^{-\beta E}\ .$$

Comme S et $T^{-1}E$ (grandeurs extensives proportionnelles à N) sont très grands devant $k_B \sim 10^{-23}$, la probabilité $p(E)$ est très "piquée" autour de la valeur $E = U$ (énergie interne macroscopique des sous systèmes) qui rend $S(E,N) - \frac{E}{T}$ maximum. Donc l'état d'équilibre observé est celui qui rend le potentiel thermodynamique $F = E - TS$ minimum. Dans la pratique il est très rare qu'on puisse déterminer exactement ce minimum. On appelle **théorie de champ moyen** une approximation qui consiste à exprimer S et E en fonction d'un paramètre d'ordre x (le taux d'aimantation d'un site dans le cas du ferromagnétisme ; *cf.* section 5.1.4), puis à calculer la valeur $x^*(T)$ de x qui minimise $F(x) = E(x) - TS(x)$ (*cf.* section 7.4.2).

▷ Généralisations ; statistiques quantiques

Si l'état A caractérise les particules contenues dans un sous volume V d'un système isolé $\mathcal{S}$, il faut préciser à la fois l'énergie E_A et le nombre de particules N_A. En utilisant la formule générale $\mathrm{d}S = T(\mathrm{d}U + P\,\mathrm{d}V - \mu\,\mathrm{d}N)$, on obtient comme précédemment la loi de probabilité dans le **formalisme grand canonique** :

$$p_A = \frac{\exp\left(k_B^{-1}\, S(U_0 - E_A, V_0 - V, N_0 - N_A)\right)}{\exp\left(k_B^{-1}\, S(U_0, V_0, N_0)\right)} = e^{-\beta PV}\, e^{-\beta(E_A - \mu N_A)}\ .$$

Ce formalisme est utile quand on considère des particules quantiques pour lesquelles les règles d'occupation des états particulaires α dépendent de la statistique (Fermi-Dirac ou Bose-Einstein). Dans ce cas, si les particules sont sans interaction, un état A est entièrement défini par l'ensemble $\{n_\alpha\}$ des nombres d'occupation des états α avec les conditions

$$\sum_\alpha n_\alpha = N_A \quad \text{et} \quad \sum_\alpha n_\alpha \epsilon_\alpha = E_A$$

($n_\alpha = 0$ ou 1 pour les **fermions** et $0, 1, 2\cdots$ pour les **bosons**). Ceci entraîne que la probabilité p_A factorise : $p_A \propto \Pi_\alpha\, e^{-\beta(\epsilon_\alpha - \mu)\, n_\alpha}$. De la loi de probabilité marginale relative au nombre d'occupation n_α d'un état $p(n_\alpha) \propto e^{-\beta(\epsilon_\alpha - \mu)\, n_\alpha}$ (loi exponentielle pour n_α), on déduit que le nombre moyen d'occupation est

$$< n_\alpha >= \sum_{n_\alpha = 0}^{1 \text{ ou } \infty} n_\alpha\, p(n_\alpha) = \frac{1}{e^{\beta(\epsilon_\alpha - \mu)} \pm 1} \qquad (+1\ \textbf{fermions},\ -1\textbf{bosons})\ ,$$

et que la variance vaut $< n_\alpha^2 > - < n_\alpha >^2 = < n_\alpha > \left(1 \mp < n_\alpha >\right)$ (utiliser $\sum_0^\infty n^p x^n = (x\partial_x)^p (1-x)^{-1}$ pour les bosons). Pour des fermions l'écart type vaut 0 si $< n_\alpha >= 1$ (état α occupé) ; pour des bosons il vaut $< n_\alpha >$ si $< n_\alpha >$ est grand. Si l'état α est peu occupé ($< n_\alpha > \ll 1$) l'écart type vaut $< n_\alpha >^{\frac{1}{2}}$ pour les bosons comme pour les fermions ; on est alors dans les conditions de la **physique statistique classique**.

REMARQUES. Le **potentiel chimique** μ est déterminé par la condition $\Sigma_\alpha < n_\alpha >= < N > \simeq N$ où N est le nombre de particules du système étudié (grandeur macroscopique dont les fluctuations sont négligeables). Dans le cas particulier des photons N n'est pas fixé et $\mu = 0$; $< n >= (e^{\beta h\nu} - 1)^{-1}$. Pour un volume V et une bande de fréquences $[\nu, \nu + \mathrm{d}\nu]$, le nombre d'états de photons est (compte tenu des polarisations) $2\,\frac{V 4\pi p^2\,\mathrm{d}p}{h^3}$ avec $p = \frac{h\nu}{c}$; on en déduit que la densité volumique spectrale d'énergie est $u(\nu, \beta) = h\nu\,\frac{8\pi\nu^2}{c^3}\,(e^{\beta h\nu} - 1)^{-1}$ (**loi de Planck**).

▷ Réponse linéaire et fluctuations (exemples)

Pour un système macroscopique paramagnétique, les états A en présence d'un champ magnétique extérieur B ont une énergie $E_A - M_A B$. La fonction de partition

$$Z(\beta, B) = \sum_A \exp -\beta(E_A - M_A B)$$

(dont l'expression précise n'importe pas) permet de calculer l'aimantation moyenne

$$< M >= \sum_A p_A M_A \qquad \left(p_A = Z^{-1}\, \exp\, -\beta\,(E_A - M_A B)\right)$$

et la variance $< (\delta M)^2 >= \sum_A p_A M_A^2 - < M >^2$:

$$< M >= \beta^{-1}\, \partial_B \ln Z \quad ; \quad < (\delta M)^2 >= \beta^{-2} \left(\frac{\partial_B^2 Z}{Z} - \left(\frac{\partial_B Z}{Z} \right)^2 \right) .$$

On en déduit la relation $\partial_B < M >= \beta < (\delta M)^2 >$: la réponse du système à une petite variation de B est une petite variation de M proportionnelle à ses fluctuations. ($\chi = \mu_0\, \partial_B < M > |_{B=0}$ est la susceptibilité.)
De même la réponse à une petite variation de température ΔT est une petite variation d'énergie interne $\Delta U = c_v\, \Delta T$ du système et on vérifie que la capacité thermique $c_v = \partial_T U$ est reliée aux fluctuations de l'énergie par :

$$T c_v = T\, \partial_T U = -\beta\, \partial_\beta U = \beta \Big(\frac{\partial_\beta^2 Z}{Z} - \big(\frac{\partial_\beta Z}{Z} \big)^2 \Big) = \beta < (\delta E)^2 >$$

(car $U =< E >= -\partial_\beta \ln Z$).

REMARQUE. A la température de la transition entre paramagnétisme et ferromagnétisme, l'expérience montre que χ devient infini (*cf.* aussi section 5.1.4). Donc $< (\delta M)^2 >= \sum_{ij} (< \mu_i \mu_j > - < \mu_i >< \mu_j >)$ (car $M = \sum_i \mu_i$) tend aussi vers l'infini. Ceci implique, réflexion laissée au lecteur, qu'il y a entre les moments individuels μ_i des corrélations à longue portée (même si les interactions sont à courte portée).

10.2.4 Estimation et lois de χ^2 (khi-deux)

▷ Estimation

Un problème fréquemment rencontré est celui de l'estimation de la valeur a d'une grandeur (non aléatoire) à partir des résultats $x_1, x_2 \cdots x_n$ de n mesures (entachées d'erreurs non systématiques) relatives à une grandeur x fonction connue $x(a)$ de a. Ici les n mesures sont indépendantes et peuvent correspondre à des protocoles expérimentaux différents, ne pas avoir le même degré de précision, *etc.* Pour prendre en compte les erreurs, l'idée est de considérer pour chaque i le nombre x_i comme l'une des valeurs possibles d'une v.a. X_i, les X_i étant n v.a. indépendantes. Un **estimateur** $\widehat{A}$ de a est une v.a. fonction des X_i qui, pour être bon, doit être **sans biais** ($< \widehat{A} >= a$), de variance la plus petite possible, et convergent ($\widehat{A}$ sûr pour $n \to \infty$). Sa détermination par la **méthode du maximum de vraisemblance** consiste à se donner une densité de probabilité

$$\Pi_i\, p_i(x_i - x(a))$$

"raisonnable" pour les erreurs $X_i - x(a)$, et à retenir la valeur $\hat{a}$ (fonction des x_i) qui maximise cette probabilité. $\hat{a}(x_1, x_2 \cdots x_n)$ est une réalisation expérimentale de l'estimateur $\widehat{A} = \hat{a}(X_1, X_2 \cdots X_n)$.

▷ Estimations conduisant à un χ^2

On se limite à deux exemples d'estimation amenant à considérer des réalisations χ^2_{exp} d'une v.a. χ^2, qui est la somme des carrés de v.a. gaussiennes centrées indépendantes.

EXEMPLE 1. Si la grandeur mesurée est directement celle qu'on veut estimer ($x \equiv a$) et si les erreurs sont gaussiennes de variances σ_i^2, alors $\Pi_i\, p_i(x_i - x) \propto \Pi_i \exp - \dfrac{(x_i - x)^2}{2\sigma_i^2}$.

Maximiser cette fonction par rapport à x revient à minimiser la quantité

$$\chi^2(x) = \sum_{i=1}^{n} \frac{(x_i - x)^2}{\sigma_i^2} \quad (\chi^2 \text{ à } \boldsymbol{n} \textbf{ degrés de libertés}) .$$

L'estimateur est $\widehat{X} = \dfrac{\sum_i w_i X_i}{\sum_i w_i}$ avec $w_i = \sigma_i^{-2}$, et sa réalisation est :

$$\hat{x} = \frac{\sum_i w_i x_i}{\sum_i w_i} .$$

Il donne un poids plus grand aux mesures les plus précises ; ce n'est que si toutes sont faites avec la même précision que l'estimation expérimentale est la moyenne arithmétique des résultats de mesure. Le χ^2 expérimental a pour expression $\chi^2_{\text{exp}} = \sum_{i=1}^{n} \dfrac{(x_i - \hat{x})^2}{\sigma_i^2}$.

EXEMPLE 2. **Estimation d'une v.a.** X. On cherche alors à estimer sa moyenne, sa variance, sa loi de probabilité, *etc.* à partir des valeurs $x_1, x_2 \cdots x_n$ obtenues lors de n mesures de X (les mesures indépendantes étant ici supposées se faire sans erreur). Comme précédemment pour chaque i, mais parce que X est aléatoire et non à cause d'erreur de mesure, x_i est une valeur possible d'une v.a. $X_i \equiv X$, les X_i étant indépendantes. On vérifie en calculant explicitement leurs moyennes que

$$\widehat{M} = \tfrac{X_1 + X_2 + \cdots + X_n}{n} \text{ et } \widehat{V} = \tfrac{1}{n-1} \textstyle\sum_{i=1}^{n} (X_i - \widehat{M})^2$$

sont des estimateurs sans biais de la moyenne et de la variance de X. Si on s'intéresse à la loi de probabilité de X (supposée discrète pour simplifier), et si lors d'une série de n mesures la valeur x_α du spectre de X a été trouvée n_α fois, $\widehat{P}_\alpha = \frac{N_\alpha}{n}$ est un estimateur sans biais de la probabilité théorique p_α de cette valeur ; N_α est la v.a. dont les valeurs n_α dépendent de la série de n mesures effectuée. On peut montrer que pour n grand les v.a. $\frac{N_\alpha - np_\alpha}{\sqrt{np_\alpha}}$ sont proches de gaussiennes centrées, réduites ; on pose alors $\chi^2_{\text{exp}} = \sum_{\alpha=1}^{s} \frac{(n_\alpha - np_\alpha)^2}{np_\alpha}$. (Ici le nombre de degrés de liberté est $r = s - 1$ car les N_α sont reliés par la contrainte $\sum_\alpha N_\alpha = n$.)

▷ Loi et test de χ^2

Après avoir obtenu une estimation se pose toujours le problème du "degré de confiance" dans cette estimation qui n'est qu'une réalisation de l'estimateur. C'est là qu'intervient la **loi du χ^2 à r degrés de liberté** : loi de la v.a. $Z = \sum_{i=1}^{r} G_i^2$ où les v.a. G_i sont des gaussiennes centrées réduites indépendantes. Les tables de χ^2 donne la valeur χ^2_p telle que $P(\{Z < \chi^2_p\}) = p$ pour $p = 0,5,\ 0,90,\ 0,95 \ldots$ Le test du χ^2 consiste alors à se donner une valeur de p (proche de 1) à laquelle correspond une valeur χ^2_p, et à comparer cette dernière au χ^2_{exp} ; le test est négatif si $\chi^2_{\text{exp}} > \chi^2_p$ (puisque la probabilité $1 - p$ très petite de cette éventualité devrait pratiquement l'exclure). Si $\chi^2_{\text{exp}} < \chi^2_p$ le test est positif, ce qui ne constitue pas une preuve de validité. Ce n'est qu'en satisfaisant simultanément plusieurs types de tests qu'une estimation devient de plus en plus plausible ; par exemple, il faut vérifier aussi que $< X^4 > - 3 < X^2 >^2 \simeq 0$ pour une v.a. gaussienne centrée.

REMARQUE. Une autre application importante est l'ajustement de données expérimentales par la méthode des moindres carrés (*cf.* section 11.3.3).

10.3 PROCESSUS ALÉATOIRES

Considérons une particule brownienne dans un fluide située en O à $t = 0$. Lors d'une expérience ω_1 son mouvement, dû aux chocs avec les molécules du fluide, est décrit par une fonction "ordinaire" $\vec{r}_1(t)$; dans une autre expérience ω_2 la trajectoire $\vec{r}_2(t)$ est différente. Ces fonctions sont des réalisations d'une fonction aléatoire $\overrightarrow{R}(t)$. En général on ne s'intéresse pas à la trajectoire entière, mais on cherche à savoir si il existe des corrélations entre les positions à des instants différents, comment se comporte le déplacement $\overrightarrow{R}(t_2) - \overrightarrow{R}(t_1)$, *etc.* Ce sont donc les v.a. $\overrightarrow{R}(t_1)$, $\overrightarrow{R}(t_2) \cdots \overrightarrow{R}(t_n)$ ($t_1, t_2 \ldots t_n$ étant des paramètres) qu'on étudie, et en pratique on se limite aux familles à un et deux instants qui caractérisent en général assez bien le processus. On commence ici par considérer la marche au hasard qui est plus simple que le mouvement brownien et peut servir d'introduction aux autres processus étudiés par la suite.

10.3.1 Marche aléatoire ; processus de diffusion ; mouvement brownien

▷ Modélisations de la diffusion ; bruit blanc

Considérons un marcheur qui, partant de l'origine O à $t = 0$, se déplace sur une droite en effectuant aux instants $0, \tau, 2\tau \cdots n\tau \cdots$ des pas indépendants $\pm l$ avec la même probabilité (de transition) $\dfrac{1}{2}$. Sa position $X(t)$ est une somme de v.a. indépendantes, centrées et de variance l^2 d'où :

$$< X(t) >= 0 \quad ; \quad < (X(t))^2 >= nl^2 = 2Dt \ \ (t = n\tau) \ .$$

D est le **coefficient de diffusion** défini par $2D = \dfrac{l^2}{\tau} = Vl$. Ces résultats restent valables quelle que soit la loi des pas pourvu qu'elle soit de moyenne nulle et de variance l^2 ; l et V sont des libre parcours moyen et vitesse typiques.

Un processus de diffusion correspond à la limite continue où τ et l tendent vers zéro, D restant fixé. Il peut aussi être considéré comme une "vision à grande échelle" de la marche aléatoire. Si $\tau = \epsilon$ le nombre de pas $n = \dfrac{t}{\epsilon}$ effectués dans l'intervalle $[0, t]$ tend vers l'infini lorsque $\epsilon \to 0$, et (*cf.* section 10.2.1) $X(t)$ devient une v.a. gaussienne. Sa variance étant $2Dt$ et sa moyenne nulle, sa densité s'écrit :

$$P_X(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp -\frac{x^2}{4Dt} \ ;$$

elle vérifie l'**E.D.P. de diffusion** (*cf.* section 8.2.2) $\partial_t P = D\, \partial_x^2 P$ pour $t > 0$. La densité de probabilité conditionnelle de $X(t_2)$ sachant que $X(t_1) = x_1$ ($t_2 > t_1 > 0$) s'appelle **probabilité de transition** du processus notée $P_{t_2,t_1}(x_2 \,/\, x_1)$; son expression

$$P_{t_2,t_1}(x_2 \,/\, x_1) = \big(4\pi D(t_2 - t_1)\big)^{-\frac{1}{2}} \exp -\frac{(x_2 - x_1)^2}{4D(t_2 - t_1)}$$

se déduit de celle de $P_X(x,t)$ en remplaçant l'intervalle $[0, t]$ par $[t_1, t_2]$ et x par $(x_2 - x_1)$. Plus généralement les déplacements $X(t_2) - X(t_1)$ sont des v.a. gaussiennes centrées, de

variance $2D(t_2 - t_1)$, et qui sont indépendantes si les intervalles $[t_1, t_2]$ ne se recouvrent pas.

REMARQUE. Si $t_2 - t_1 = \mathrm{d}t$ est infinitésimal on peut poser $X(t_2) - X(t_1) = \mathrm{d}X(t)$. Les physiciens appellent **bruit blanc** (gaussien) le "processus limite" défini par :

$$B(t) = \lim_{\mathrm{d}t \to 0} \frac{\mathrm{d}X(t)}{\mathrm{d}t} \ .$$

Il vérifie $< B(t) >= 0$ et $< B(t)\, B(t') >= 0$ pour $t \neq t'$ (indépendance des déplacements $\mathrm{d}X(t)$ et $\mathrm{d}X(t')$). Comme la variance de $(t_2 - t_1)^{-1}\ (X(t_2) - X(t_1))$ vaut $2D(t_2 - t_1)^{-1}$ on voit que $< B(t)\, B(t') >$ est singulier pour $t = t'$. On justifie l'expression $< B(t)\, B(t') >= 2D\, \delta(t - t')$ en remarquant que cela redonne bien :

$$< \big(X(t)\big)^2 >= \int_0^t \int_0^t < B(t')\, B(t'') > \, \mathrm{d}t'\, \mathrm{d}t'' = 2Dt \ .$$

Autre modélisation. Considérons une chaîne d'atomes et un électron pouvant sauter d'un atome vers les deux atomes voisins distants de a avec une égale probabilité $\lambda\, \mathrm{d}t$ dans l'intervalle $[t, t + \mathrm{d}t]$. Alors si $P(x,t)$ est sa probabilité de présence à l'abscisse x à l'instant t, on a (*cf.* $p(A) = \sum_i p(A/B_i)\, p(B_i)$)

$$P(x, t + \mathrm{d}t) = \big(P(x + a, t) + P(x - a, t)\big)\, \lambda\, \mathrm{d}t + P(x,t)\, (1 - 2\lambda\, \mathrm{d}t) \ ,$$

ce qui donne l'équation maîtresse $\partial_t P(x,t) = \big(P(x+a,t) + P(x-a,t) - 2P(x,t)\big)\, \lambda$. Dans la limite continue $a \to 0$, $\lambda \to \infty$ et $\lambda a^2 = D$, on retrouve l'équation $\partial_t P = D\, \partial_x^2 P$.

▷ E.D.P. de Fokker-Planck

Ce modèle se généralise en considérant des probabilités de transition par unité de temps de x vers $x \pm a$ asymétriques et inhomogènes de la forme $\lambda \pm \epsilon(x)$. Ceci conduit aux termes supplémentaires $-P(x + a, t)\, \epsilon(x + a) + P(x - a, t)\, \epsilon(x - a)$ et, dans la limite continue avec $A(x) = 2a\epsilon(x)$ fixé, à l'équation :

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\, (A(x)\, P(x,t)) + D\, \frac{\partial^2 P(x,t)}{\partial x^2} \ .$$

$A(x)$ est la **vitesse de dérive** associée au déplacement moyen

$$a(\lambda + \epsilon(x))\, \mathrm{d}t - a(\lambda - \epsilon(x))\, \mathrm{d}t = A(x)\, \mathrm{d}t$$

dans le temps $\mathrm{d}t$. Si $A = 0$ l'équation décrit de la diffusion pure. Si $D = 0$ l'équation $\partial_t P(x,t) + \partial_x\, (A(x) P(x,t)) = 0$ décrit l'évolution temporelle d'une densité de points sur une droite lorsque $\dot{x} = A(x)$ (système déterministe ; *cf.* section 8.2.3).

▷ Equation de Langevin à une dimension : $m\dot{V} + \alpha V = F(t)$

Elle décrit le **mouvement brownien** d'une particule dans un fluide en supposant que la force due aux chocs avec les molécules du fluide se sépare en un effet moyen de friction αV et une force $F(t)$ de moyenne nulle, indépendante de la position $X(t)$ et de la vitesse $V(t)$ de la particule et de type bruit blanc gaussien. Pour un petit intervalle de temps Δt, la variation de vitesse ΔV contient une partie "dérive" $-\alpha m^{-1}\, V\, \Delta t$ (qui empêche la vitesse de devenir trop grande) et une partie "diffusion" telle que $< (\Delta V)^2 >= m^{-2} < \left(\int_0^{\Delta t} F(t')\, \mathrm{d}t'\right)^2 >= 2D_V\ \Delta t$ en appelant D_V le coefficient de diffusion dans l'espace des vitesses ; (si α était nul $< V^2(t) >$ tendrait vers l'infini avec t). La densité $P(v,t)$ obéit donc à l'équation :

$$\partial_t P = \partial_v \left(\tfrac{\alpha}{m}\, v\, P\right) + D_V\, \partial_v^2 P \ .$$

Sa valeur d'équilibre $P_{\rm eq} \propto \exp \frac{-\alpha v^2}{2mD_V}$ (telle que $\partial_t P_{\rm eq} = 0$) correspond à la **loi de Boltzmann** si $m^2 D_V = \alpha k_B T$. Il existe donc une relation entre les fluctuations de $F(t)$ (force due aux chocs pour $V = 0$), le coefficient de friction α (caractérisant la réponse linéaire des molécules à un mouvement de la particule) et la température T. On obtient la diffusion en position X en écrivant :

$$m^{-1}\alpha < XV >= m^{-1} < XF > - < X\dot{V} >= - < X\dot{V} >=< V^2 > - \tfrac{\rm d}{{\rm d}t} < XV > \ .$$

On en déduit que pour $t \gg \frac{m}{\alpha}$ on a $\alpha < XV > \simeq m < V^2 >= k_B T$ et, comme $\frac{\rm d}{{\rm d}t} < X^2 >= 2 < XV >$, il vient :

$$< X^2(t) > \simeq 2Dt \ \text{ avec } \ D = \tfrac{k_B T}{\alpha} (\textbf{relation d'Einstein}) \ .$$

10.3.2 Processus de Markov ; probabilités de transition ; bilan détaillé

▷ Définition

Ce sont des processus aléatoires tels que la connaissance de la valeur de $\underline{X}(t)$ à un instant t_0 résume tout le passé, c.a.d. détermine la loi de probabilité de $\underline{X}(t)$ pour $t > t_0$. Le processus de diffusion (pour $X(t)$), de Langevin (pour $V(t)$), de même que la marche au hasard en sont des exemples (avec respectivement t continu et t discret). Si les valeurs prises par le processus sont discrètes, celui-ci est déterminé par les **probabilités de transition** $P_{t,t_0}(\underline{x}/\underline{x}_0)$ donnant la probabilité (conditionnelle) pour que $\underline{X}(t) = \underline{x}$ sachant que $\underline{X}(t_0) = \underline{x}_0$; elles deviennent des densités si les valeurs $\underline{x}$ sont continues. Comme on peut passer de $\underline{x}_0$ à $\underline{x}$ en transitant par une valeur quelconque $\underline{y}$ à un instant intermédiaire τ, ces probabilités vérifient

$$P_{t,t_0}(\underline{x}/\underline{x}_0) = \sum_{\underline{y}} P_{t,\tau}(\underline{x}/\underline{y})\, P_{\tau,t_0}(\underline{y}/\underline{x}_0)$$

(relation analogue à celle $F_{t,t_0} = F_{t,\tau} \circ F_{\tau,t_0}$ écrite à la section 6.1.1. pour les systèmes dynamiques, mais ici avec $\tau \in [t_0, t]$).

▷ Chaînes de Markov ; bilan détaillé

Il s'agit de processus à temps discret ($t_n = n\tau$) où le système peut transiter d'un état j vers un état i avec la probabilité constante P_{ij} ($\sum_i P_{ij} = 1$). Si le nombre d'états est fini, la probabilité $p_i(n+1)$ d'être dans l'état i à l'instant $(n+1)\tau$ est reliée aux probabilités $p_j(n)$ d'être dans un état j à l'instant précédent $n\tau$ par :

$$p_i(n+1) = \sum_j P_{ij}\, p_j(n) \qquad \text{ou sous forme matricielle} \quad \underline{p}\,(n+1) = \mathbf{P}\,\underline{p}\,(n) \ .$$

L'exemple le plus simple de **matrice de transition** est $\mathbf{P} = \begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix}$ (avec $a, b \in [0,1]$) ; elle correspond à un système à deux états.

Propriétés des matrices P.
1) Les valeurs propres vérifient $|\lambda| \leq 1$; en effet, si $\underline{v}$ est un vecteur propre on a $|\lambda|\,|v_i| = |\sum_j P_{ij}\, v_j| \leq \sum_j P_{ij}\,|v_j|$, et en sommant sur i on obtient le résultat.

2) Si la matrice est symétrique les valeurs propres sont réelles. La valeur propre -1 est exceptionnelle (pour la matrice 2×2 ci-dessus il faudrait $a = b = 0$, cas d'une transition certaine). La valeur propre 1 correspond au vecteur propre dont les composantes sont égales ; elle est en général non dégénérée. En effet, de $v_i = \sum_j P_{ij}\, v_j = \sum_j P_{ji}\, v_i$, on déduit $\sum_j P_{ij}(v_j - v_i) = 0$ pour tout i (car $\sum_j P_{ji} = 1$ et $P_{ji} = P_{ij}$) ; ceci implique l'égalité des composantes car si il en existait une, v_i, plus petite que les autres, l'égalité ci-dessus serait violée (en supposant que les P_{ij} ne sont pas nuls).
3) Sauf exception, la matrice $\mathbf{P}^{-1}$ n'est pas une matrice de transition ; (pour la matrice 2×2 ci-dessus il faudrait soit $a = b = 1$ soit $a = b = 0$). Les processus de Markov décrivent en général des **phénomènes irréversibles**.

Microréversibilité et distribution microcanonique. En mécanique quantique les probabilités de transition entre états sont symétriques (*cf.* section 4.4.2) : $P_{ij} = P_{ji}$. Si on considère pour $i, j \cdots$ les états microscopiques d'un système macroscopique isolé (point de vue microcanonique), on déduit des propriétés de la matrice $\mathbf{P}$ qu'il y a une seule distribution d'équilibre qui est la distribution uniforme où tous les états sont équiprobables. Toute distribution initiale tend vers elle. En effet en décomposant l'état initial $\underline{p}(0)$ sur les vecteurs propres de $\mathbf{P}$, on voit que, au cours de l'évolution, seule sa composante sur le vecteur propre $\underline{p}^{\rm eq}$ correspondant à la valeur propre $\lambda = 1$ va subsister : $\underline{p}(n) = \mathbf{P}^n\, \underline{p}(0) \to \underline{p}^{\rm eq}$ quand $n \to \infty$ (car les autres valeurs propres vérifient $|\lambda| < 1$).

On dit qu'une **distribution d'équilibre** (définie par $\mathbf{P}\underline{p}^{\rm eq} = \underline{p}^{\rm eq}$) obéit à la relation de **bilan détaillé** si il y a autant de transitions de i vers j que de j vers i : $P_{ji}\, p_i^{\rm eq} = P_{ij}\, p_j^{\rm eq}$ (*cf.* exemple ci-dessous et section 11.5). Cette relation qui correspond à l'absence de courant dans l'espace des états (figure 11a) n'est pas a priori nécessaire pour $\underline{p}^{\rm eq}$ (figure 11b).

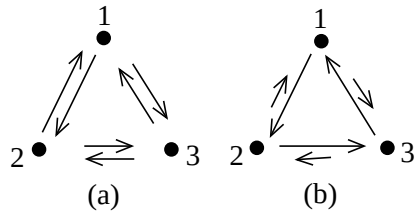


FIGURE 11

▷ Equation maîtresse

Si le temps est continu, on introduit des probabilités de transition par unité de temps w_{ij}. L'**équation maîtresse** satisfaite par les probabilités de présence $p_i(t)$ s'obtient intuitivement en faisant le“ bilan des arrivées et des départs” :

$$\dot{p}_i(t) = \sum_j w_{ij}\, p_j(t) - \sum_j w_{ji}\, p_i(t) \ .$$

On peut la considérer comme limite continue de l'équation $p_i(n+1) = \sum_j P_{ij}\, p_j(n)$ d'une chaîne de Markov. Un état d'équilibre correspond à $\underline{\dot{p}}^{\rm eq} = 0$ et la condition de bilan détaillé s'écrit $w_{ji}\, p_i^{\rm eq} = w_{ij}\, p_j^{\rm eq}$.

EXEMPLE : **coefficients d'Einstein**. Pour décrire un atome à deux niveaux (le niveau 2 étant le niveau excité) en équilibre avec le rayonnement thermique à la température T ($\beta = (k_B T)^{-1}$), Einstein a posé $w_{12} = A(\nu) + B_{12}(\nu)\, u(\nu,\beta)$ et $w_{21} = B_{21}(\nu)\, u(\nu,\beta)$ où $u(\nu,\beta)$ est la densité volumique spectrale d'énergie, B_{12} et A sont respectivement associés aux émissions induite et spontanée et B_{21} à l'absorption (*cf.* cours de physique). En écrivant la relation de bilan détaillé $w_{12}\, e^{-\beta\epsilon_2} = w_{21}\, e^{-\beta\epsilon_1}$ (avec $\epsilon_2 - \epsilon_1 = h\nu$), et en comparant $u(\nu,\beta) = \dfrac{A(\nu)}{B_{21}(\nu)\, e^{\beta h\nu} - B_{12}(\nu)}$ avec la loi de Planck (*cf.* section 10.2.3), il en a déduit l'égalité $B_{12} = B_{21}$ et l'expression du rapport $\dfrac{A}{B_{12}}$.

Remarque. Le bilan détaillé est en fait une conséquence de la relation $\frac{w_{12}}{w_{21}} = \frac{<n>+1}{<n>}$ déduite de la quantification du champ électromagnétique (*cf.* section 4.4.3), et dans laquelle le nombre n de photons a été remplacé par sa moyenne $< n >= (e^{\beta h\nu} - 1)^{-1}$ (*cf.* section 10.2.3).

10.3.3 Processus stationnaires ; théorème de Wiener-Khintchine ; ergodicité

▷ Définitions

Un processus est **stationnaire** si ses propriétés statistiques ne dépendent pas du choix de l'origine des temps. Pour le décrire on se limite en général à sa moyenne (constante, prise égale à zéro dans la suite) et sa **fonction de corrélation** définie par :

$$\Gamma(\tau) =< \overline{X}(t)\, X(t+\tau) > \ .$$

Elle vérifie $\Gamma(\tau) = \overline{\Gamma}(-\tau)$ (changement $t \to t-\tau$) et $\Gamma(0) \geq |\Gamma(\tau)|$ (conséquence de $< |X(t+\tau) + \lambda\, X(t)|^2 > \geq 0$ quel que soit λ). En général $\Gamma(\tau)$ fait intervenir un temps de corrélation (ou de cohérence) τ_c au delà duquel $X(t)$ et $X(t+\tau)$ ne sont quasiment plus corrélés ; τ_c est nul pour un bruit blanc car alors $\Gamma(\tau) \propto \delta(\tau)$ (*cf.* section 10.3.1). Parmi les processus stationnaires fréquemment rencontrés figurent les **processus gaussiens** pour lesquels toute famille de v.a. $X(t_1), X(t_2) \cdots X(t_n)$ est gaussienne, et qui sont donc entièrement caractérisés par $\Gamma(\tau)$, et les **processus poissonniens** tels que $X(t)$ change de manière aléatoire aux instants T_i définis à la section 10.2.2 .

EXEMPLE 1. Signal lumineux émis par un atome soumis à des chocs dans un gaz (figure 10b) :

$$X(t) = a\, \exp -i(2\pi\nu_0 t + \Phi(t))$$

où $\Phi(t)$, v.a. équirépartie sur $[0, 2\pi]$, change aux instants T_i. Si entre t_1 et t_2 aucun choc ne s'est produit (évènement A), on a $\Phi(t_2) = \Phi(t_1)$; si au moins un choc s'est produit dans cet intervalle (évènement $\overline{A}$), $\Phi(t_1)$ et $\Phi(t_2)$ sont indépendantes. En écrivant

$$\begin{aligned} < \overline{X_{t_1}}\, X_{t_2} > &=< \overline{X_{t_1}}\, X_{t_2}/A > \ P(A) + < \overline{X_{t_1}}\, X_{t_2}/\overline{A} > \ P(\overline{A}) \\ &= |a|^2\, \exp -2i\pi\nu_0(t_2 - t_1)\, P(A) + 0 \times P(\overline{A}) \end{aligned}$$

avec $P(A) = \exp -\lambda|t_2 - t_1|$ (*cf.* section 10.1.2), on obtient la fonction de corrélation :

$$\Gamma(\tau) =< \overline{X_{t_1}}\, X_{t_2} >= |a|^2\, \exp -2i\pi\nu_0\, \tau\, \exp -\lambda|\tau| \ .$$

EXEMPLE 2. La vitesse du **mouvement brownien**, donnée par $V(t) = \int_{-\infty}^{t} \exp -\frac{\alpha}{m}(t-s)\, \frac{F(s)}{m}\, \mathrm{d}s$, est une v.a. gaussienne car somme de v.a. gaussiennes. De $< F(s)\, F(s') >= 2D_V m^2\, \delta(s-s')$, on déduit (petit calcul) $< V(t_1)\, V(t_2) >= D_V \frac{m}{\alpha}\, \exp -\frac{\alpha}{m}|\tau| = \Gamma(\tau)$, avec $\tau = t_2 - t_1$ (*cf.* aussi section 5.3.3).

▷ Théorème de Wiener-Khintchine

La transformée de Fourier (tronquée) $\widehat{X}_T(\nu) = \displaystyle\int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-i2\pi\nu t}\, X(t)\, \mathrm{d}t$ vérifie pour $T \gg \tau_c$:

$$\frac{1}{T} < \widehat{X}_T(\nu')\, \overline{\widehat{X}}_T(\nu) > = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-i2\pi\nu\tau}\, \Gamma(\tau)\, \mathrm{d}\tau \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-i2\pi(\nu'-\nu)t}\, \mathrm{d}t$$
$$\simeq \widehat{\Gamma}(\nu)\ \mathrm{sinc}\pi(\nu'-\nu)T\ .$$

DÉMONSTRATION : on fait le changement de variable $t' = t + \tau$ dans $\int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{i2\pi\nu t}\, e^{-i2\pi\nu' t'} < \overline{X}(t)\, X(t') >\, \mathrm{d}t\, \mathrm{d}t'$.

Pour $\nu = \nu'$ le sinus cardinal vaut 1 et on obtient le théorème :

$$\widehat{\Gamma}(\nu) = \lim_{T\to\infty} T^{-1} < |\widehat{X}_T(\nu)|^2 >\ .$$

Ce résultat est semblable à celui établi à la section 5.3.3 pour la densité spectrale de puissance d'un signal chaotique stationnaire. Pour l'exemple 1 ci-dessus $\widehat{\Gamma}(\nu) \propto \left(4\pi^2(\nu-\nu_0)^2 + \lambda^2\right)^{-1}$: les chocs conduisent à un **profil de raie lorentzien**.

Pour $\nu \neq \nu'$ et T grand le sinus cardinal vaut zéro ; on interprète ce résultat comme l'absence de corrélation entre les composantes de Fourier d'un signal stationnaire.

▷ Ergodicité

Une question importante dans la pratique est de savoir si les moyennes statistiques $< \cdots >$ relatives à un processus stationnaire $X(t)$ peuvent être remplacées par les moyennes temporelles $\lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (\cdots)\, \mathrm{d}t$ relatives à *une* réalisation "typique" $x(t) = X(t,\omega)$ du processus. Un processus pour lequel ceci est possible est dit **ergodique** et on doit avoir en particulier :

$< X(t) >= \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t')\, \mathrm{d}t'$; $< \overline{X}(t) X(t+\tau) >= \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \overline{x}(t')\, x(t'+\tau)\, \mathrm{d}t'$.

Pour que, par exemple, l'égalité relative à la moyenne soit satisfaite il est nécessaire que, dans la limite $T \to \infty$, la v.a. $Y = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t)\, \mathrm{d}t$ devienne sure, c.a.d. que σ_Y tende vers zéro ; de $\sigma_Y^2 = \frac{1}{T^2} \times \int_{-\frac{T}{2}}^{\frac{T}{2}} \int_{-\frac{T}{2}}^{\frac{T}{2}} < \overline{X}(t) X(t') >\, \mathrm{d}t\, \mathrm{d}t' = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \Gamma(\tau)\, \mathrm{d}\tau$ (calcul identique au précédent), on déduit que ceci est satisfait notamment si $\Gamma(\tau) \to 0$ quand $\tau \to \infty$ (cas des exemples ci-dessus). Plus généralement les "**processus à perte de mémoire**", tels que les v.a. $X(t)$ et $X(t')$ deviennent indépendantes pour $|t'-t|$ fini, sont des exemples de processus ergodiques.

Chapitre **11**

Analyse numérique ; physique discrète

L'étude numérique d'un problème physique se déroule schématiquement en trois étapes. La première concerne sa formalisation mathématique ; les dix premiers chapitres de ce livre y ont été consacrés, l'accent étant mis sur l'importance des concepts mathématiques. La seconde concerne l'élaboration d'un modèle numérique "sensé mimer au mieux" le modèle mathématique ; elle relève de l'analyse numérique abordée ici en raison de son importance de plus en plus grande en physique. Enfin la dernière concerne la mise en œuvre sur l'ordinateur de la méthode précédente ; c'est la phase de programmation, située hors du champ de cet ouvrage.

Une fois formalisé, par exemple sous la forme d'une équation différentielle (E.D.), un problème physique n'est pas pour autant résolu. Même si la solution analytique de l'équation est connue, le physicien est intéressé non seulement par les valeurs prises par cette solution en fonction de la variable, mais aussi en fonction des conditions initiales (C.I.) (ou aux limites (C.L.)), des paramètres de l'équation ... Se posent alors de nombreux problèmes numériques, tel que celui du choix, crucial, du pas de discrétisation (de la variable, des paramètres ...) qui conditionne la pertinence et la précision d'un calcul, ou tel que l'interpolation pour le tracé de graphes.

Mais très souvent le problème physique n'a pas de solution explicite, par exemple s'il conduit à une E.D. linéaire à coefficients variables ou à une E.D. non linéaire. Le recours au numérique est alors indispensable ; il implique, pour ces exemples, la discrétisation des opérations de dérivation et donc la transformation d'un système dynamique (S.D.) continu en un S.D. discret. La vérification que ce dernier constitue une bonne approximation du S.D. continu est un problème important et non trivial, souvent facilité par la connaissance préalable (physique et mathématique) des propriétés qualitatives de la solution. Sous réserve de cette vérification la simulation du problème physique permet une "expérimentation numérique" y compris dans des domaines de paramètres non accessibles à l'expérimentation physique. Inversement, l'étude a priori de modèles numériques peut aider à la compréhension et à la formalisation mathématique de problèmes physiques.

Les méthodes d'analyse numérique ne concernent évidemment pas que les E.D. (ou les E.D.P.). La physique linéaire, l'analyse des signaux, la formulation de solutions physiques en terme d'extremum, ainsi que la physique statistique y font largement appel. Inversement, les méthodes numériques peuvent s'inspirer des outils théoriques développés dans ces secteurs (*cf.* ci-après, section 11.2.3, la méthode de Lax dans laquelle on ajoute des termes diffusifs à certains schémas numériques pour les rendre stables ou, section 11.5.2, la méthode du recuit simulé qui utilise la loi de Boltzmann pour trouver le minimum principal d'une fonction en évitant de rester "piégé" dans un minimum secondaire).

Compte tenu de l'étendue du sujet nous nous contentons dans ce chapitre de décrire quelques méthodes numériques parmi les plus utiles aux physiciens, en précisant leurs limites (stabilité, précision ...) et en donnant une idée de leurs applications.

11.1 DISCRÉTISATION

11.1.1 Représentation des nombres ; erreurs ; stabilité numérique

▷ Représentation des nombres

Dans un ordinateur les nombres sont représentés en binaire (suite de 0 et de 1) dans un format dont la longueur dépend du nombre de bits utilisés (32 en simple précision et 64 en double précision). Dans la représentation à point flottant un nombre s'écrit

$$\pm M\, 2^{e-q} \quad \text{avec} \quad \frac{1}{2} \leq M < 1 \ .$$

M est la mantisse, e l'exposant et q un nombre fixe (décallage d'exposant).

$$\underbrace{\times}_{\text{signe}} \quad \underbrace{\times\times\times\times\times\times\times\times}_{\text{exposant } e} \quad \underbrace{(1)\times}_{\text{mantisse } M}$$

En simple précision le premier des 32 bits spécifie le signe (0 pour + et 1 pour -) ; les huit bits suivants codent l'exposant e compris entre 0 et $2^8 - 1 = 255$, l'exposant total $e - q$ variant de -126 à 129 ($q = 126$) ; enfin les 23 bits restants (en fait 24 car la condition $\frac{1}{2} \leq M$ entraîne que le premier bit de M est toujours égal à 1 et est sous-entendu dans cette représentation) codent la mantisse comprise entre $\frac{1}{2}$ et $1 - 2^{-24}$. Il en résulte que le domaine des nombres x représentables est

$$2^{-127} \simeq 5,9\,10^{-39} \leq x \leq 2^{129}(1 - 2^{-24}) \simeq 6,8\,10^{38} \ .$$

Mais il est important de remarquer que tous les nombres de ce domaine ne peuvent pas être représentés exactement (par exemple $\frac{1}{3}$ dont la mantisse devrait être la suite périodique infinie 101010...).

▷ Nombres en physique (rappel)

De très grands et de très petits nombres peuvent apparaître en physique. Cependant les valeurs des grandeurs physiques sont toujours relatives à un système d'unités. Avant de rentrer des valeurs dans l'ordinateur il est donc important d'effectuer une analyse dimensionnelle et d'ordre de grandeur du problème afin de mettre en évidence des nombres

sans dimension, d'éventuelles propriétés d'échelle (*cf.* section 1.3.2)... puis de choisir des unités adaptées. (Par exemple $\hbar \simeq 10^{-34}\, J.s$ et $m \simeq 10^{-30}\, kg$ disparaissent de l'E.D.P. de Schrödinger pour un électron dans un potentiel $V(\vec{r})$, de valeur typique V_0, si les unités de temps et de longueur sont prises égales à $\hbar V_0^{-1}$ et $\hbar\,(2mV_0)^{-\frac{1}{2}}$.)

▷ Erreurs d'arrondi et de troncature

Le fait que la mantisse ne comporte qu'un nombre fini de bits entraîne des erreurs d'arrondi dans la représentation des nombres et lors des opérations (additions, multiplications...). En simple précision, les deux plus petits nombres positifs représentables $\frac{1}{2}2^{-q}$ et $\left(\frac{1}{2}+2^{-24}\right)2^{-q}$ sont séparés de $2\delta x = 2^{-24}2^{-q}$. Il en résulte une erreur relative d'arrondi

$$\epsilon = \frac{\delta x}{x} = 2^{-24} = 6\,.\,10^{-8}\ .$$

(En double précision le plus petit nombre représentable est 10^{-308} et l'erreur relative d'arrondi vaut 10^{-16}.)

Une autre source d'erreur, à priori indépendante de la précédente, est l'erreur de troncature liée à l'utilisation de schémas avec un pas h de discrétisation fini. Regardons sur un exemple comment les deux types d'erreurs interviennent sur la précision d'un résultat.

EXEMPLE. Si on calcule la dérivée d'une fonction par

$$f'(x) = h^{-1}\big(f(x+h) - f(x)\big) \quad \text{ou} \quad f'(x) = (2h)^{-1}\big(f(x+h) - f(x-h)\big)\ ,$$

l'erreur de troncature, estimée par un développement de Taylor, se comporte comme $h\,|f''|$ (ordre 1) ou $h^2\,|f'''|$ (ordre 2) ; elle diminue avec h. L'ordre de grandeur de l'erreur d'arrondi est dans les deux cas $\epsilon \times h^{-1}|f|$ (étant entendu que h est toujours représenté exactement) ; elle croit avec h^{-1}. La somme est minimale pour une valeur de h optimale $h_{\text{opt}} = \left[\epsilon\frac{|f|}{|f''|}\right]^{\frac{1}{2}}$ ou $h_{\text{opt}} = \left[\epsilon\frac{|f|}{|f'''|}\right]^{\frac{1}{3}}$ et vaut $\sqrt{\epsilon}\,\Delta$ ou $\epsilon^{\frac{1}{3}}\,\Delta$, où $\Delta \simeq \frac{|f|}{|f'|} \simeq \frac{|f'|}{|f''|} \simeq \frac{|f''|}{|f'''|}$ est une longueur typique sur laquelle f et ses dérivées varient (*cf.* chap. 5, figure 3). Pour cette valeur optimale l'erreur relative sur f' se comporte donc comme $\sqrt{\epsilon}$ ou $\epsilon^{\frac{2}{3}}$. On remarque qu'elle peut être nettement plus grande que ϵ (par exemple de quatre ordres de grandeur si on compare $\sqrt{\epsilon}$ à ϵ pour $\epsilon = 10^{-8}$). On remarque aussi que cette erreur s'approche de ϵ quand l'ordre de l'erreur de troncature (puissance en h) augmente. On verra comment augmenter cet ordre tout en partant d'un schéma numérique d'ordre bas (extrapolation de Richardson).

▷ Stabilité d'un schéma numérique

Un autre problème relatif aux erreurs, auquel il faut veiller, concerne leur évolution dans un schéma récurrent. Celui-ci peut les amplifier, auquel cas il est dit instable. Ceci n'est pas grave si le schéma numérique est précis et le nombre d'itérations limité, mais est en général à éviter. Un exemple élémentaire (généralisé plus loin aux E.D. et aux E.D.P.) est le calcul des puissances $u_n = x^n$ du nombre $x = \frac{\sqrt{5}-1}{2}$. Comme $x^2 + x - 1 = 0$ on peut utiliser la récurrence $u_{n+1} = u_{n-1} - u_n$ avec $u_0 = 1$ et $u_1 = 0,680... = x$ (à la précision de la machine). Une erreur se propage alors selon la règle $\delta u_{n+1} = \delta u_{n-1} - \delta u_n$

qui admet pour solution générale $\delta u_n = \alpha x^n + \beta(-x)^{-n}$ (*cf.* section 8.1.1). L'erreur est donc amplifiée à chaque pas : elle évolue comme x^{-n} avec $x < 1$. En simple précision δu_n devient de l'ordre de u_n vers $n \simeq 16$ (alors que $x^{16} \simeq 10^{-4}$).

11.1.2 Dérivation et intégration ; extrapolation de Richardson

▷ Dérivation

On considère des fonctions $f(t)$, une discrétisation de pas $t_{n+1} - t_n = h$ constant et on pose $f(t_n) = f_n$. Les approximations les plus simples (dites "**avant**", "**arrière**" et "**centrée**") de la **dérivée première** $f'(t_n)$ sont :

$$f'_{n+} = \frac{f_{n+1} - f_n}{h} \quad , \quad f'_{n-} = \frac{f_n - f_{n-1}}{h} \quad , \quad f'_{n0} = \frac{f_{n+1} - f_{n-1}}{2h} \ .$$

Elles correspondent aux pentes des segments BC, AB et AC sur la figure 1 et sont respectivement d'ordre 1, 1 et 2 (erreurs en h, h et h^2) ; par exemple $|f'_{n+} - f'(t_n)| \leq \frac{h}{2} \sup_t |f''(t)|$, avec $t \in [t_n, t_{n+1}]$. La première approximation de la **dérivée seconde** $f''(t_n)$ est $f''_n = \dfrac{f_{n+1} - 2f_n + f_{n-1}}{h^2}$ (*cf.* sections 6.3 et 8.1.2).

REMARQUE. En dimension 2, le **laplacien** $\big(\partial_x^2 + \partial_y^2\big) f(x,y)$ d'une fonction $f(x,y)$, discrétisée selon $f(x_0 + jh, y_0 + kh) = f_{j,k}$, est représenté par :

$$\Delta f_{j,k} = h^{-2}\big(f_{j+1,k} + f_{j-1,k} + f_{j,k+1} + f_{j,k-1} - 4f_{j,k}\big) \ .$$

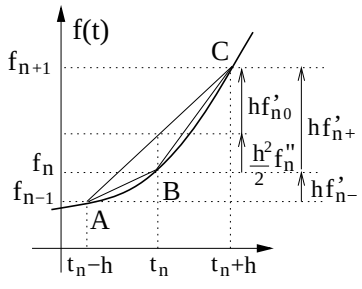


FIGURE 1

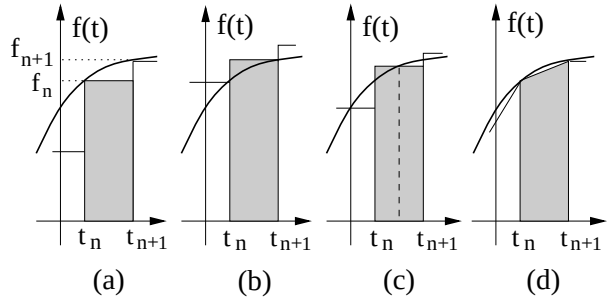


FIGURE 2

▷ Intégration

L'intégrale de $f(t)$ entre t_n et t_{n+1} peut, par exemple, être approchée par :

$$h\, f_n \quad , \quad h\, f_{n+1} \quad , \quad h\, f\Big(\frac{t_n + t_{n+1}}{2}\Big) \quad , \quad h\, \frac{f_n + f_{n+1}}{2} \ .$$

Ces approximations, qui sont respectivement dites méthode du **rectangle à gauche**, du **rectangle à droite**, du **point milieu** et du **trapèze**, sont illustrées sur la figure 2.

Les erreurs correspondantes sont bornées par $h^2 \sup|f''|$ pour (a) et (b), $\frac{h^3}{4} \sup|f'''|$ pour (c) et $\frac{h^3}{12} \sup|f'''|$ pour (d). Pour un intervalle fini Δt composé de N sous intervalles de longueur $h = \frac{\Delta t}{N}$, les erreurs dépendent donc de N comme N^{-1} ($N^{-2} \times N$) pour les méthodes des rectangles et N^{-2} ($N^{-3} \times N$) pour celles du point milieu et du trapèze. Par exemple pour le trapèze :

$$S = \int_{t=t_0}^{t+\Delta t=t_N} f(t)\,\mathrm{d}t = \frac{\Delta t}{N}\Big(\frac{f_0}{2} + f_1 + \cdots + f_{N-1} + \frac{f_N}{2}\Big) + \mathcal{O}\Big(\frac{1}{N^2}\Big)\ .$$

L'intérêt de cette méthode est qu'on peut montrer que l'erreur est une fontion paire de h, donc du type $AN^{-2} + BN^{-4} + \cdots$ en fonction du nombre de pas. Donc si S_N est la valeur approchée de l'intégrale pour N pas (N pair), et $S_{N/2}$ celle pour $\frac{N}{2}$ pas, on a $S_N = S + AN^{-2} + \cdots$, $S_{N/2} = S + 4AN^{-2} + \cdots$ et :

$$\frac{4}{3}S_N - \frac{1}{3}S_{N/2} = S + \mathcal{O}\Big(\frac{1}{N^4}\Big)\ .$$

Donc, sans diminuer h mais en tenant compte de la valeur approchée pour $2h$, on a amélioré l'estimation de S de 2 ordres ! Cette technique d'amélioration de la précision, basée sur la dépendance en N (donc en h) de S_N, est un cas particulier de l'**approche de Richardson de la valeur limite** ; cette technique, qui peut s'appliquer à de nombreux problèmes, consiste à utiliser les résultats obtenus pour différents pas (h, $2h$,...) pour approcher, par extrapolation, la valeur exacte associée à la valeur limite $h = 0$.

11.1.3 Nombres aléatoires ; méthode de Monte-Carlo

▷ Générateurs de nombres aléatoires

Dans les ordinateurs, les entiers aléatoires répartis de façon **équiprobable** entre 0 et un grand entier $m - 1$, sont engendrés (pour les générateurs les plus courants) en utilisant la méthode de la congruence linéaire. Les entiers $I_1, I_2, \cdots$ sont alors obtenus, à partir d'un entier quelconque I_0 (le **germe**), par la récurrence :

$$I_{n+1} = aI_n + c \quad \text{modulo } m\ .$$

(Il existe des valeurs optimum pour a et c qui dépendent de celle de m.) Une telle suite étant périodique de période m, on voudrait pouvoir prendre pour m le plus grand nombre représentable par la machine mais la multiplication par a conduit à des "overflows". (En simple précision on ne peut pas dépasser $m \simeq 10^6$, nombre "petit", mais l'utilisation d'un générateur faisant intervenir successivement plusieurs germes différents permet d'allonger la période.)

On engendre les lois courantes de probabilité en considérant des fonctions des I_n. Pour la **loi uniforme** sur $[0, 1]$, on utilise la suite $x_n = \frac{I_n}{m}$. Pour la **loi exponentielle**, on utilise la suite $y_n = -\lambda^{-1} \ln(x_n)$; en effet, de $p(y)\,\mathrm{d}y = p(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|\mathrm{d}y$ (*cf.* section 10.1.2) on déduit $p(y) = \lambda \exp[-\lambda y]$. Pour la **loi gaussienne** de moyenne μ et de variance σ^2, on utilise la suite $y_n = \mu + \sqrt{-2\sigma^2 \ln x_n}\cos 2\pi x_{n+1}$.

DÉMONSTRATION. Soient x et x' deux v.a. équiréparties sur $[0,1]$ et $y = \sqrt{-2\sigma^2 \ln x'}\cos 2\pi x$, $y' = \sqrt{-2\sigma^2 \ln x'}\sin 2\pi x$ (alors $x = \frac{1}{2\pi}\arctan\frac{y'}{y}$ et $x' = \exp\left[-\frac{1}{2\sigma^2}(y^2+y'^2)\right]$). De $p(y,y')\,\mathrm{d}y\,\mathrm{d}y' = p(x,x')\left|\frac{\partial(x,x')}{\partial(y,y')}\right|\mathrm{d}y\,\mathrm{d}y'$, on déduit $p(y,y') = \left|\frac{\partial(x,x')}{\partial(y,y')}\right| = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}y^2\right]\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}y'^2\right]$. Donc y et y' sont deux v.a. gaussiennes indépendantes de moyenne nulle et de variance σ^2 ; il suffit alors de poser $x' = x_n$, $x = x_{n+1}$ et $y + \mu = y_n$.

Pour une loi quelconque $p(x)$, une méthode générale est la méthode de rejet (*cf.* remarque ci-dessous).

▷ Intégration par la méthode de Monte-Carlo

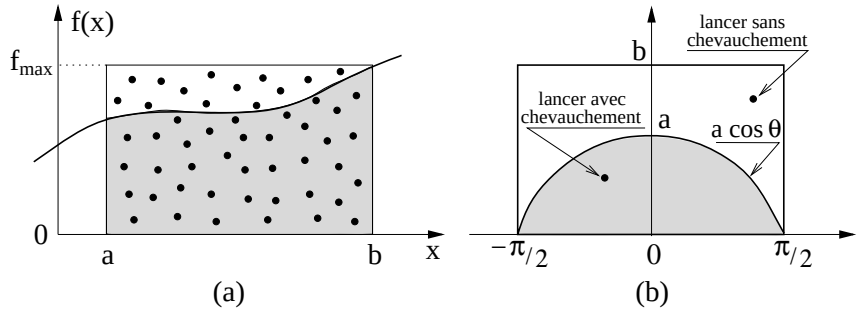


FIGURE 3

Le calcul de $I = \int_a^b f(x)\,\mathrm{d}x$ $(f(x) > 0)$ par la méthode de Monte-Carlo consiste à tirer N points au hasard à l'intérieur d'un rectangle d'aire $A = (b-a)f_{\max}$ qui contient le graphe ; les coordonnées x_i et y_i des points sont des v.a. équiréparties sur $[a,b]$ et $[0, f_{\max}]$ (figure 3a). I est alors estimée par l'aire A multipliée par la fraction de points qui sont "tombés" en dessous de la courbe ; l'erreur se comporte comme $N^{-\frac{1}{2}}$.

REMARQUE. **Méthode de rejet**. Si $f(x)$ est une loi de probabilité sur $[a,b]$, les coordonnées x_i des points conservés sont distribuées selon cette loi.

EXEMPLE historique : aiguille de **Buffon**. La détermination "expérimentale" de π en jetant N fois "au hasard" une aiguille de longueur a sur un parquet dont les lames ont une largeur $b > a$, et en comptant la proportion des cas où l'aiguille chevauche deux lames, revient à calculer l'aire sous la courbe représentée sur la figure 3b par la méthode de Monte-Carlo (réflexion laissée au lecteur qui montrera que la probabilité de chevauchement vaut $\frac{2a}{\pi b}$).

La méthode d'intégration de Monte-Carlo se généralise à d dimensions. En dimension 1 ce n'est pas la plus performante, même dans des versions plus élaborées que celle présentée ci-dessus (Monte-Carlo adaptatif qui utilise des échantillons aléatoires non uniformément répartis par exemple). Par contre c'est la plus efficace pour le calcul d'intégrales multiples lorsque le domaine d'intégration a une forme compliquée ou (et) en grande dimension.

11.2 RÉSOLUTION NUMÉRIQUE D'E.D. ET D'E.D.P.

11.2.1 Systèmes dynamiques ; schémas d'Euler et de Runge Kutta

Rappelons que toute E.D. ou système d'E.D. d'ordre quelconque se ramène à un système d'ordre 1 pour n fonctions (*cf.* section 6.1). Pour simplifier on considère le cas $n = 1$:

$$\dot{x}(t) = f(x(t),t) \quad , \quad x(0) = x_0 \ .$$

Le temps étant supposé discrétisé de façon régulière dans un certain domaine, on pose :

$$x_n = x(t_n) \quad , \quad f_n = f(x(t_n), t_n) \quad \text{avec} \quad t_{n+1} - t_n = h \ .$$

La résolution numérique du S.D. consiste à exprimer x_{n+1} en fonction des valeurs connues x_{n-k} $(k \geq 0)$.

▷ Schémas principaux

Les méthodes les plus simples sont les schémas d'**Euler**

$$x_{n+1} = x_n + hf_n \quad \text{ou} \quad x_{n+1} = x_n + hf_{n+1}$$

qui correspondent à l'intégration de $f(x(t), t)$ par les méthodes du rectangle. Ce sont des schémas d'ordre 1 (une schéma d'intégration étant d'ordre n si l'erreur tend vers 0 comme h^{n+1}). Le premier est **explicite** (x_{n+1} n'apparaît pas dans le membre de droite). Le second est **implicite** et demande en général la résolution d'une équation à chaque pas (par la méthode de Newton-Raphson par exemple). Les deux ont été combinés pour l'obtention graphique de trajectoires solutions de $\ddot{\overrightarrow{r}} = \overrightarrow{F}(\overrightarrow{r})$ à partir de $\overrightarrow{r}_{n+1} = \overrightarrow{r}_n + h\overrightarrow{v}_n$, $\overrightarrow{v}_n = \overrightarrow{v}_{n-1} + h\overrightarrow{F}_n$ (*cf.* section 6.3, figures 16 où $h = 1$). Un autre schéma implicite est celui de **Crank-Nicolson**

$$x_{n+1} = x_n + \frac{h}{2}(f_n + f_{n+1}) \quad \text{(ordre 2)}$$

qui correspond à l'intégration de f par la méthode du trapèze.

Schémas implicites et explicites sont souvent combinés dans des méthodes dites de **prédicteur-correcteur**. Ainsi celui d'Euler explicite $x_{n+1} = x_n + hf_n$ peut servir à fournir une valeur de x_{n+1} pour le membre de droite de l'équation de Cranck-Nicholson. Si on se limite à la première itération on obtient la formule de **Runge-Kutta d'ordre 2** (R.K.2)

$$x_{n+1} = x_n + \tfrac{h}{2}\big(f(x_n, t_n) + f(x_n + hf_n, t_n + h)\big)$$

équivalente à $x_{n+1} = x_n + hf\big(x_n + \frac{1}{2}hf_n, t_n + \frac{h}{2}\big)$. Le schéma R.K.2 donne donc x_{n+1} en estimant $\dot{x}$ au point milieu, au lieu des points extrêmes comme les schémas d'Euler. (On notera que x_{n+1} obtenu par le schéma R.K.2 s'écrit aussi $x^R_{n+1} = \frac{1}{2}(x^E_{n+1} + x^E_{n+2})$ où x^E_{n+1} et x^E_{n+2} sont obtenus à partir de x_n par le schéma d'Euler explicite ; *cf.* figure 4 qui correspond à $\dot{x} = -x$ et $h = \frac{1}{2}$.)

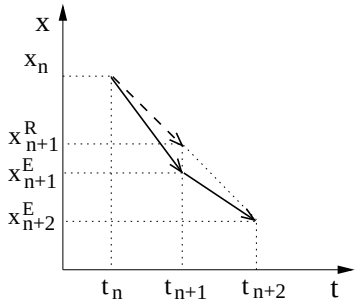


FIGURE 4

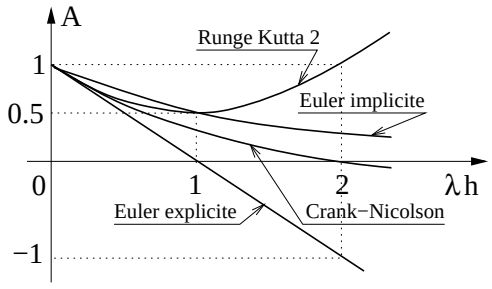


FIGURE 5

Un schéma très utilisé parce que très précis est celui de **Runge-Kutta d'ordre 4** (erreur en h^5) donné par $x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$ avec $k_1 = hf(x_n, t_n)$, $k_2 = hf\Big(x_n + \frac{k_1}{2}, t_n + \frac{h}{2}\Big)$, $k_3 = hf\Big(x_n + \frac{k_2}{2}, t_n + \frac{h}{2}\Big)$ et $k_4 = hf(x_n + k_3, t_n + h)$. Il

estime $\dot{x}$ en t_n, $t_n + h$ et deux fois en $t_n + \frac{h}{2}$.
Le lecteur vérifiera qu'appliqués au cas (trivial) $\dot{x} = -\lambda x$, les schémas d'Euler explicite, R.K.2 et R.K.4 donnent respectivement $x(t+h) = (1-\lambda h)\,x(t)$, $x(t+h) = \big(1 - \lambda h + \frac{\lambda^2 h^2}{2}\big)\,x(t)$ et $x(t+h) = \big(1 - \lambda h + \frac{\lambda^2 h^2}{2} - \frac{\lambda^3 h^3}{6} + \frac{\lambda^4 h^4}{24}\big)\,x(t)$ au lieu de $x(t+h) = e^{-\lambda h}x(t)$.

▷ Stabilité et précision

L'étude de la **stabilité** de ces schémas consiste à estimer l'effet sur x_{n+1} d'une perturbation sur x_n. Un schéma est stable si δx_n n'est pas amplifié ; cette condition s'écrit :

$$|A| \leq 1 \quad \text{avec} \quad A = \frac{\delta x_{n+1}}{\delta x_n} = \frac{\partial x_{n+1}}{\partial x_n}\ .$$

Par exemple pour l'équation $\dot{x}(t) = -\lambda x(t) + g(t)$ on obtient $\delta x_{n+1} = (1-\lambda h)\,\delta x_n$ pour Euler explicite. Ce schéma est donc stable si $\lambda h \leq 2$. En calculant de même A pour Euler implicite, Crank-Nicolson et R.K.2 le lecteur vérifiera que les schémas implicites sont plus stables que les schémas explicites (figure 5). Un exemple d'instabilité est donnée par le schéma **leapfrog** $x_{n+1} = x_{n-1} - 2h(\lambda x_n + g_n)$ (dérivée centrée) qui conduit à $A^2 + 2h\lambda\,A - 1 = 0$, soit à deux valeurs pour A dont une vérifie $|A| > 1$ quel que soit h.

Précision. La condition de stabilité est nécessaire au calcul mais elle ne garantit pas sa précision. Pour celà on sait qu'il faut que h soit nettement plus petit que le temps caractéristique τ d'évolution de $x(t)$ (l'erreur étant d'ordre $(h\tau^{-1})^{n+1}$ pour un schéma d'ordre n). Par exemple pour $\dot{x} = -\lambda x$, $\tau = \lambda^{-1}$ et il faut $\lambda h \ll 1$. Par contre, pour $\dot{x}(t) = -\lambda x(t) + g(t)$, avec g lentement variable "à l'échelle λ^{-1}" (temps caractéristique $\tau_g = K\lambda^{-1}$ avec $K \gg 1$), il suffit que $\lambda h \ll K$ (en dehors du régime transitoire). Dans ce dernier cas on a intérêt à choisir un schéma stable pour des valeurs de λh grandes devant 1 afin d'éviter des temps de calcul trop longs et l'accumulation d'erreurs.

Une façon de déterminer h pour obtenir une précision donnée, quand celle-ci ne peut pas être estimée a priori, est la suivante. Soit $x^*(t+2h)$ la solution exacte déduite de $x(t)$ et soient x_{2h} et x_{h+h} les valeurs correspondantes données par le schéma numérique pour un pas $2h$ (en une étape) et pour deux itérations de pas h. Pour le schéma R.K.4 par exemple, les erreurs sont $x_{2h} - x^*(t+2h) = C\,(2h)^5$ et $x_{h+h} - x^*(t+2h) = 2C\,h^5$. La différence $x_{2h} - x_{h+h} \simeq C\,(2h)^5$ permet d'estimer C et donc l'erreur $\Delta = C\,h^5$ à chaque pas h. On peut alors ajuster le pas h en cours de calcul pour atteindre la précision requise : si Δ_0 est cette précision il faut prendre un nouveau pas $h_0 = h\,\big(\frac{\Delta_0}{\Delta}\big)^{\frac{1}{5}}$.

11.2.2 E.D. avec conditions aux limites ; méthode de tir

Considérons (*cf.* section 6.4) une E.D. linéaire du second ordre pour une fonction $y(x)$ représentant l'amplitude d'une **onde stationnaire** de pulsation ω. L'onde est définie sur un intervalle $[a, b]$ aux extrémités duquel sont imposées des C.L. impliquant linéairement y et sa dérivée. L'E.D. du second ordre se ramène à deux E.D. du premier ordre

$$\frac{\mathrm{d}y_i(x)}{\mathrm{d}x} = \sum_j a_{ij}(x,\omega)\,y_j(x) \quad (i,j=1,2)$$

(avec par exemple $y_1 = y$ et $y_2 = y'$). Celles-ci peuvent être discrétisées sous la forme

$$y_i^{n+1} - y_i^n = \frac{1}{2}(x_{n+1} - x_n)\sum_j (a_{ij}^n y_j^n + a_{ij}^{n+1} y_j^{n+1}) \quad \text{(schéma d'ordre 2)}$$

avec $y_j^n = y_j(x_n)$, $a_{ij}^n = a_{ij}(x_n, \omega)$ $(n = 1, 2, \cdots N-1)$. Ce système de $2(N-1)$ équations linéaires s'écrit matriciellement $\underline{y}^{n+1} = \mathbf{A}_n \underline{y}^n$ où $\underline{y}^n$ désigne le vecteur colonne de composantes (y_1^n, y_2^n) et $\mathbf{A}_n$ une matrice 2×2 jouant le rôle de matrice de transfert discrète. Les C.L. s'écrivent $\alpha_1 y_1(a) + \alpha_2 y_2(a) = 0$ et $\beta_1 y_1(b) + \beta_2 y_2(b) = 0$.

Dans une **méthode de tir** (figure 6a où $y(a) = y(b) = 0$) le principe consiste à se donner (à un facteur λ près) $\underline{y}(a)$ satisfaisant la C.L. en a puis à calculer, comme dans un problème de C.I., $\underline{y}(b)$ par application de la matrice de transfert, la constante λ étant déterminée par une normalisation sur $\underline{y}(b)$ (par exemple $y_2(b) = 1$ si la C.L. en b le permet). On calcule alors la quantité $\Delta(\omega) = \beta_1 y_1(b) + \beta_2 y_2(b)$, fonction de la pulsation ω qui doit être nulle pour la solution. Avec une discrétisation de ω, on repère les intervalles où $\Delta(\omega)$ change de signe puis on applique dans chacun de ces intervalles une méthode de recherche de zéro (*cf.* section 11.3.2) pour obtenir les pulsations propres. Au lieu de a, la méthode de tir peut partir de b ou de a et b simultanément (avec recollement des deux solutions en un point intermédiaire).

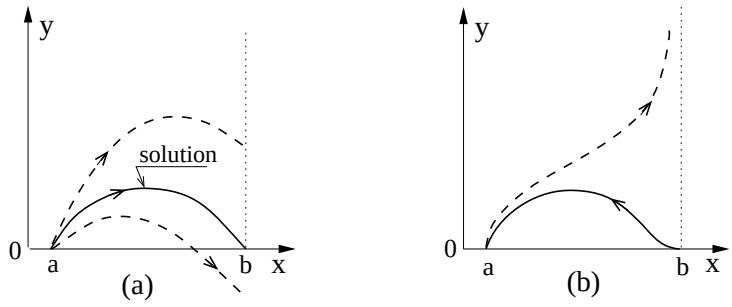


FIGURE 6

Le choix entre ces diverses possibilités devient déterminant si l'E.D. du second ordre est singulière en un point de $[a, b]$. Par exemple si la singularité est en b, il existe près de $x = b$ deux solutions puissance du type $(b-x)^{\gamma_+}$ et $(b-x)^{\gamma_-}$ avec $\gamma_+ > 0$ et $\gamma_- < 0$. ω n'étant pas égal à une pulsation propre au départ, un tir à partir de a va avoir une composante sur la solution divergente et conduire automatiquement à des valeurs infinies pour $y(x)$ en b (figure 6b). Il faut donc partir de b avec la condition $y'(b) = -\gamma_+ y(b)$. (On a vu à la section 8.3.2 qu'en symétrie sphérique une E.D.P. faisant intervenir le laplacien conduit, après séparation des variables, à une E.D. singulière en $r = 0$ pour la partie radiale avec des solutions en r^l et $r^{-(l+1)}$.)

11.2.3 E.D.P. avec conditions initiales : propagation, diffusion

▷ E.D.P. de propagation

Considérons l'équation de propagation la plus simple :

$$\partial_t f + c\, \partial_x f = 0 \quad (c \text{ constante}) \ .$$

Soient Δt et Δx les pas de discrétisation et f_j^n la valeur de f pour $t = t_n$ et $x = x_j$. Un schéma numérique a priori naturel (Euler explicite pour t, dérivée centrée pour x) est :

$$(a) \qquad f_j^{n+1} = f_j^n - c\,\Delta t\, \frac{f_{j+1}^n - f_{j-1}^n}{2\,\Delta x} = f_j^n - \frac{h}{2}\left(f_{j+1}^n - f_{j-1}^n\right) \quad \text{avec} \quad h = \frac{c\,\Delta t}{\Delta x} \ .$$

(On observera que Δt et Δx n'apparaîssent qu'à travers h qui est sans dimension.) L'étude de sa stabilité consiste à voir comment évolue une perturbation δf_j^n. La relation (a) étant linéaire à coefficients constants, il suffit (*cf.* section 8.2.1) de considérer une perturbation de type onde plane de vecteur d'onde $\vec{k}$. Après un pas celle-ci est multipliée par une constante. On vérifie en effet que :

$$\delta f_j^n \propto e^{ik\,j\Delta x} \longrightarrow \delta f_j^{n+1} = (1 - ih\,\sin k\,\Delta x)\,\delta f_j^n = A_a f_j^n \ .$$

On aurait de même $A_b = 1 + h - he^{ik\,\Delta x}$ pour le schéma (b) $f_j^{n+1} = f_j^n - h(f_{j+1}^n - f_j^n)$. Etonnamment, ces deux schémas sont instables : $|A_{a,b}| > 1$. La raison est d'origine physique : l'E.D.P. décrit une propagation à la vitesse c vers les $x > 0$ et donc f_j^{n+1} ne peut dépendre a priori que des f_j^n, f_{j-1}^n, f_{j-2}^n..., mais pas de f_{j+1}^n. Les figures 7a et 7b illustrent la propagation des données f_j^n et la dépendance de f_j^{n+1} vis a vis des f_j^n.

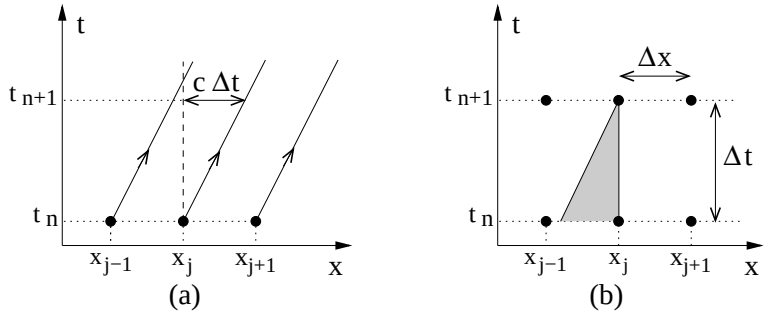


FIGURE 7

La figure 7b permet de comprendre qualitativement pourquoi le schéma

$$(c) \quad f_j^{n+1} = f_j^n - h\left(f_j^n - f_{j-1}^n\right) \quad (\text{pour lequel } A_c = 1 - h + h\,e^{-ik\,\Delta x})$$

est stable si $h < 1$: les données $f(t_n, x)$ nécessaires à l'obtention de $f(t_{n+1}, x_j)$ correspondent à des valeurs de x comprises entre x_{j-1} et x_j.

Le lecteur vérifiera que les schémas

$$(d) \quad f_j^{n+1} = \tfrac{1}{2}(f_{j+1}^n + f_{j-1}^n) - \tfrac{h}{2}\,(f_{j+1}^n - f_{j-1}^n) \quad \text{et} \quad (e) \quad f_j^{n+1} = f_j^{n-1} - h\,(f_{j+1}^n - f_{j-1}^n)$$

sont aussi stables si $h < 1$ ($A_d = \cos k\,\Delta x - ih\,\sin k\,\Delta x$ et $A_e = -ih\,\sin k\,\Delta x \pm (1 - h^2\,\sin^2 k\,\Delta x)^{\frac{1}{2}}$, conséquence de $A_e^2 = 1 - 2ihA_{(e)}\,\sin k\,\Delta x$). (d) corrige (a) par l'ajout d'un terme de diffusion en $f_{j+1}^n - 2f_j^n + f_{j-1}^n$ (**méthode de Lax**) et (e) en introduisant f_j^{n-1} au lieu de f_j^n (**schéma leapfrog**) ; dans les deux cas ces modifications font que les données $f(t_n, x)$ pour $x \in [x_j, x_{j+1}]$ deviennent nécessaires pour l'obtention de $f(t_{n+1}, x_j)$, ce qui justifie la présence de f_{j+1}^n au second membre des schémas (d) et (e). L'intérêt de (e) est que $|A_e|$ ne dépend pas de k et que la dépendance en k de la phase est la plus petite. La figure 8 représente les courbes $A(k)$ dans le plan complexe pour ces différents cas (avec $h = \frac{c\,\Delta t}{\Delta x} = \frac{1}{2}$). Il faut en pratique $h \leq \frac{1}{5}$.

REMARQUES. 1) L'analyse ci-dessus reste valable si c dépend de f, la propagation d'une valeur f_j^n se faisant alors à la vitesse $c(f_j^n)$; les difficultés numériques apparaîssent avec l'onde de choc (*cf.* chap. 7, figure 3) car elle met en jeu de grandes fréquences spatiales, *i.e.* $k \simeq (\Delta x)^{-1}$.

2) L'**E.D.P. de propagation du second ordre** se discrétise naturellement sous la forme :

$$\partial_t{}^2 f - c^2 \partial_x{}^2 f = 0 \longrightarrow f_j^{n+1} - 2f_j^n + f_j^{n-1} = h^2(f_{j+1}^n - 2f_j^n + f_{j-1}^n) \quad \text{avec } h = \tfrac{c\,\Delta t}{\Delta x} \ .$$

L'étude de stabilité analogue à celle faite ci-dessus conduit à $(A-1)^2 = -4Ah^2 \sin^2 k\frac{\Delta x}{2}$, équation qui admet deux solutions avec $A_1 A_2 = 1$. Il y a stabilité pour $h \leq 1$ (alors $|A_{1,2}| = 1$). Cette condition peut se comprendre à partir de figures analogues aux figures 7a et b, la propagation se faisant maintenant dans les deux sens $x > 0$ et $x < 0$. Pour $h = 1$ l'E.D.P. discrétisée s'écrit $f_j^{n+1} + f_j^{n-1} = f_{j+1}^n + f_{j-1}^n$ et a pour solution exacte $f_j^n = u(j+n) + v(j-n)$ comme l'équation d'onde.

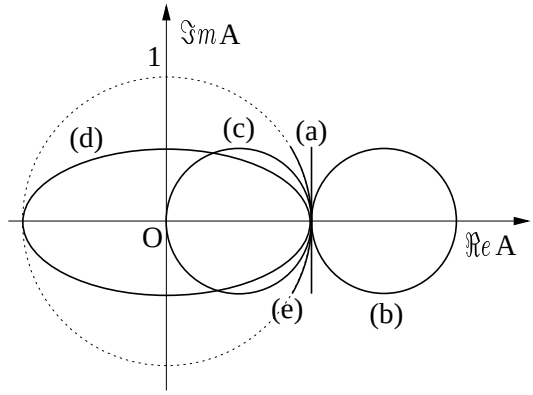


FIGURE 8

▷ E.D.P. de diffusion

L'E.D.P. de diffusion se discrétise naturellement sous la forme :

$$\partial_t f = D\partial_x{}^2 f \longrightarrow f_j^{n+1} - f_j^n = h\big(f_{j+1}^n - 2f_j^n + f_{j-1}^n\big) \quad \text{avec} \quad h = \frac{D\,\Delta t}{(\Delta x)^2} \ .$$

L'étude de stabilité conduit (petit calcul) à $A = 1 - 4h \sin^2 k\frac{\Delta x}{2}$ et à la condition $h \leq \frac{1}{2}$. L'interprétation physique de cette condition est illustrée sur les figures 9a et 9b qui montrent comment les données diffusent et comment f_j^{n+1} dépend des f_j^n.

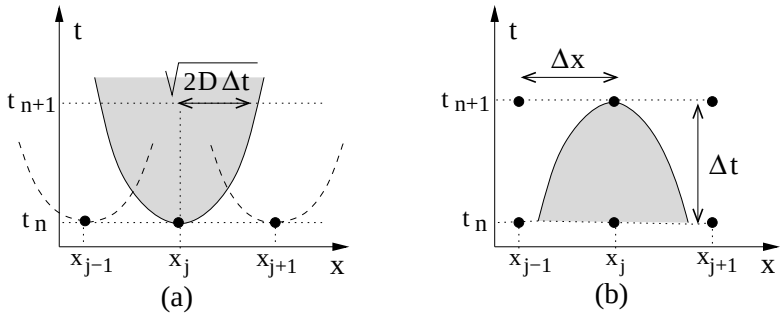


FIGURE 9

Typiquement, une diffusion sur une distance L demande un temps τ tel que $\frac{D\tau}{L^2} \sim 1$ (*cf.* section 10.3.1). Cependant, comme l'échelle de longueur L des phénomènes étudiés doit nécessairement vérifier $L \gg \Delta x$, les temps caractéristiques d'évolution $\frac{L^2}{D}$ de ces phénomènes vont impliquer un nombre de pas de temps intermédiaires $\frac{L^2}{D\,\Delta t} = \frac{L^2}{(\Delta x)^2}$ très grand. Dans la méthode de **Crank-Nicolson** (du second ordre en temps), le schéma numérique et son facteur d'amplification sont :

$$f_j^{n+1} - f_j^n = \frac{D\,\Delta t}{2(\Delta x)^2}\left[\left(f_{j+1}^{n+1} - 2f_j^{n+1} + f_{j-1}^{n+1}\right) + \left(f_{j+1}^n - 2f_j^n + f_{j-1}^n\right)\right] \quad ; \quad A = \frac{1-2h\,\sin^2 k\frac{\Delta x}{2}}{1+2h\,\sin^2 k\frac{\Delta x}{2}}\ .$$

On observe que $A < 1$ pour tout h, ce qui permet de choisir des Δt en fonction de la physique, et non de la grille Δx. L'obtention de f_j^{n+1} à partir des f_j^n est un problème linéaire faisant intervenir des matrices tridiagonales (*cf.* section 11.4.1).

REMARQUE. L'**E.D.P. de Schrödinger** $i\partial_t\Psi = H\Psi$ (avec $H = -{\partial_x}^2 + V(x)$ pour un choix convenable d'unités) est analogue à l'E.D.P. de diffusion, mais il faut respecter le caractère unitaire de l'évolution. Pour la résoudre on approxime $e^{-iH\,\Delta t}$ par le produit $e^{i\,\Delta t\,\partial_x^2}\,e^{-i\,\Delta t\,V(x)}$. Le passage de Ψ_j^n à Ψ_j^{n+1} se fait alors en deux étapes en utilisant la transformée de Fourier discrète et son inverse (*cf.* section 5.3.2) :

$$\Psi_j^n \to \widehat{\Psi}_q^{n+\frac{1}{2}} = \frac{1}{N}\sum_{j=0}^{N-1} e^{-ik_q j}\,e^{-i\,\Delta t\,V_j}\,\Psi_j^n \to \Psi_j^{n+1} = \sum_{q=0}^{N-1} e^{ik_q j}\,e^{-ik_q^2\,\Delta t}\,\widehat{\Psi}_q^{n+\frac{1}{2}} \quad \text{avec} \quad k_q = \frac{2\pi q}{N}\ .$$

11.3 APPROXIMATION DE FONCTIONS ; INTERPOLATION ; MOINDRES CARRÉS ; MÉTHODE DE BÉZIER

L'approximation de fonctions revêt des formes très variées. Un premier type de problèmes concerne des fonctions définies par des séries infinies. L'approximation consiste alors souvent à utiliser des "représentations vectorielles finies" $f(x) = \sum_{n=1}^{N} c_n\phi_n(x)$ où les $\phi_n(x)$ peuvent être des polynômes (de Tchebytchev par exemple), des polynômes par morceaux (B-splines par exemple), des fonctions sinusoïdales .... Ces représentations servent aussi au tracé de graphes, à l'intégration numérique (intégration de Gauss), à la résolution numérique d'E.D. (les c_n étant alors les inconnues à déterminer)...
Dans un autre type de problèmes on dispose d'un ensemble de données $y = y_i$ pour $x = x_i$ $(i = 1, 2, \cdots, N)$ dont on souhaite une traduction analytique $y = f(x)$. Si les y_i sont "exacts" il s'agit d'un problème d'interpolation ; l'approximation par des polynômes (de Lagrange) ou des polynômes par morceaux ("cubic-splines") passe alors par les points (x_i, y_i)). Si par contre les y_i sont entachés d'erreurs, l'interpolation exacte n'a plus de sens et on utilise une méthode de moindres carrés qui fait passer "au mieux" l'approximation au milieu du nuage de points.

11.3.1 Approximations polynomiales (Tchebytchev, B-splines...)

▷ Bases de polynômes orthogonaux

Une base $\{P_N,\ N = 0, 1, 2\cdots\}$ est caractérisée par un intervalle de définition $[a, b]$ et un produit scalaire :

$$< P, Q >= \int_a^b w(x)\,P(x)\,Q(x)\,\mathrm{d}x\ .$$

Le développement de $f(x)$ sur cette base est alors donnés par :

$$f(x) = \sum_n c_n\,P_n(x) \quad ; \quad c_n = \left(< P_n, P_n >\right)^{-\frac{1}{2}}\ < P_n, f >\ .$$

Par exemple les polynômes de **Legendre** $L_n(x)$ (*cf.* section 4.1.2) correspondent à l'intervalle $[-1, 1]$ et $w(x) = 1$, ceux de **Tchebytchev** $T_n(x)$ à $[-1, 1]$ et $w(x) = (1-x^2)^{-\frac{1}{2}}$, ceux de **Hermite** $H_n(x)$ (*cf.* section 4.4.4) à $[-\infty, \infty]$ et $w(x) = e^{-x^2}$, ceux de **Laguerre** $L_n^\alpha(x)$ à $[0, \infty]$ et $w_\alpha(x) = x^\alpha e^{-x}$... Pour chaque base P_n est de degré n, a n zéros (un dans chaque intervalle défini par ceux de P_{n-1}) et il existe des relations simples entre les

polynômes de différents degrés (par exemple $T_{n+1} = 2x\,T_n - T_{n-1}$) et entre les polynômes et leurs dérivées (par exemple $(1-x^2)\,T_n' = -nx\,T_n + n\,T_{n-1}$).

Les polynômes de **Tchebytchev**

$$T_n(x) = \cos(n \arc\cos x) \quad (= \cos n\varphi)\ ,$$

qui ont leurs $n+1$ extrema tous égaux en valeur absolue (figure 10), jouent un rôle privilégié dans l'approximation de fonctions. En effet, ils possèdent la propriété remarquable que, parmi tous les polynômes de degré n (avec un coefficient de x^n normalisé à 1), $2^{-n}T_n(x)$ est celui dont les oscillations ont la plus petite amplitude sur $[-1,1]$.

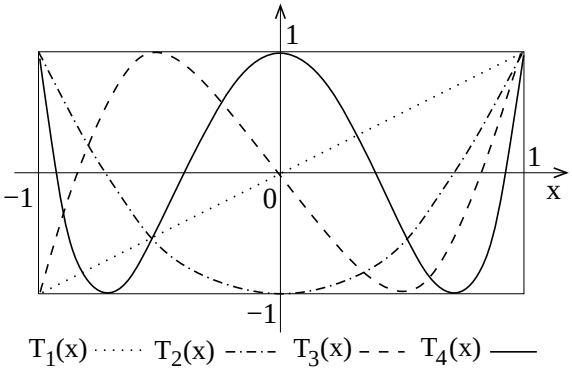


FIGURE 10

Ces oscillations étant uniformes, l'approximation d'une fonction $f(x)$ par le développement limité

$$f(x) \simeq \frac{c_0}{2} + \sum_{k=1}^{n} c_k\,T_k(x) \quad \text{avec} \quad c_k = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)\,T_k(x)}{\sqrt{1-x^2}}\,\mathrm{d}x \qquad \Big(\frac{\mathrm{d}x}{\sqrt{1-x^2}} = \mathrm{d}\varphi\Big)$$

conduit "pratiquement" au polynôme P de degré n qui minimise $\sup\limits_{x\in[-1,1]} |f(x) - P(x)|$. Comme de plus ces polynômes vérifient la relation d'orthogonalité discrète

$$\sum_{i=1}^{N} T_k(x_i)T_l(x_i) = \frac{N}{2}\delta_{kl} \text{ si } k \text{ ou } l \neq 0 \quad \text{et} \quad \sum_{i=1}^{N} T_0(x_i)T_0(x_i) = N\ ,$$

où les $x_i = \cos\Big(\dfrac{\pi}{N}(i-\dfrac{1}{2})\Big)$ sont les zéros de $T_N(x) = \cos N\varphi$ $(N > k, l)$, les coefficients c_k du développement limité peuvent être calculés par la formule exacte

$$c_k = \frac{2}{N}\sum_{i=1}^{N} f(x_i)\,T_k(x_i)\ .$$

REMARQUES. 1) En analyse on utilise beaucoup les développements limités de Taylor $\sum_{k=0}^{N} \frac{x^k}{k!} f^{(k)}(0)$. Pour eux l'erreur ϵ, d'ordre $|x|^{N+1}$, croit quand on s'éloigne du point de développement et n'est donc pas uniforme sur l'intervalle considéré (qu'on peut toujours ramener à $[-1,1]$ par un changement de variable). L'intérêt d'un développement de Tchebytchev est d'obtenir la même erreur ϵ mais uniformément sur $[-1,1]$ et avec un polynôme de degré inférieur à N.

2) Les polynômes ne constituent pas la seule façon d'approcher une fonction. On peut aussi utiliser des

fractions rationnelles. Par exemple l'**approximation de Padé** d'une série

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \longrightarrow R(x) = \frac{\sum_0^p b_i x^i}{1+\sum_1^q c_j x^j} = \frac{P(x)}{Q(x)} \ ,$$

où les b_i et les c_j sont obtenus à partir des a_k par identification des développements formels, est souvent une approximation "efficace" qui peut même représenter $f(x)$ au delà du rayon de convergence de la série initiale (*cf.* $f(x) = \sum_{k=0}^{\infty} x^k \to R(x) = (1-x)^{-1}$).

Intégration de Gauss. De façon générale, l'utilisation de polynômes orthogonaux comme base pose le problème du calcul des coefficients du développement par intégration. La méthode de Gauss, qui joue sur le choix des points d'interpolation x_i, est très efficace. Si les x_i sont les zéros du polynôme $P_N(x)$ associé au poids $w(x)$ et les w_i des constantes adaptées (à ne pas confondre avec $w(x_i)$), on montre que

$$\int_a^b w(x)\, g(x)\, \mathrm{d}x \simeq \sum_{i=1}^{N} w_i\, g(x_i) \ ;$$

cette expression est exacte si $g(x)$ est un polynôme de degré inférieur à $2N$ et constitue une bonne approximation pour g quelconque. Un exemple vient d'être vu avec $P_N = T_N$ et $w_i = \frac{2}{N}$; dans ce cas, l'erreur $\sum_{k=1}^{N} c_k\, T_k(x)$ est majorée par $\dfrac{\pi}{2^{2N-1}(2N)!} \times \sup_{x\in[-1,1]} |g^{(2N)}(x)|$. Pour $P_N = L_N$ on a $w_i = 2\big((1-x_i^2)\, L'_N(x_i)\big)^{-1}$ et l'erreur est d'ordre $((2N)!)^{-1} \sup |g^{(2N)}(x)|$.

▷ B-splines

On se donne une suite de nombres $x_1 \le x_2 \le \cdots \le x_N$ appelés noeuds. Les fonctions $B_{i,k}(x)$ d'ordre $k \ge 1$ sont alors définies par la récurrence (sur k) :

$$B_{i,k}(x) = \frac{x - x_i}{x_{i+k-1} - x_i} B_{i,k-1}(x) + \frac{x_{i+k} - x}{x_{i+k} - x_{i+1}} B_{i+1,k-1}(x) \ ,$$

avec $B_{i,1}(x)$ la fonction porte égale à 1 sur $[x_i, x_{i+1}]$. Les $B_{i,k}(x)$ sont des polynômes par morceaux de support $[x_i, x_{i+k}]$, positifs, de degré $k-1$, dérivable $k-2$ fois (pour $k \ge 3$) et tels que $\sum_i B_{i,k}(x) = 1$. La figure 11a donne des exemples pour $k = 1, 2, 3$ et des noeuds x_i régulièrement espacés (B-splines uniformes). Notons que plusieurs noeuds peuvent cependant être confondus et il est souvent utile de considérer ce cas (*cf.* ci-dessous).

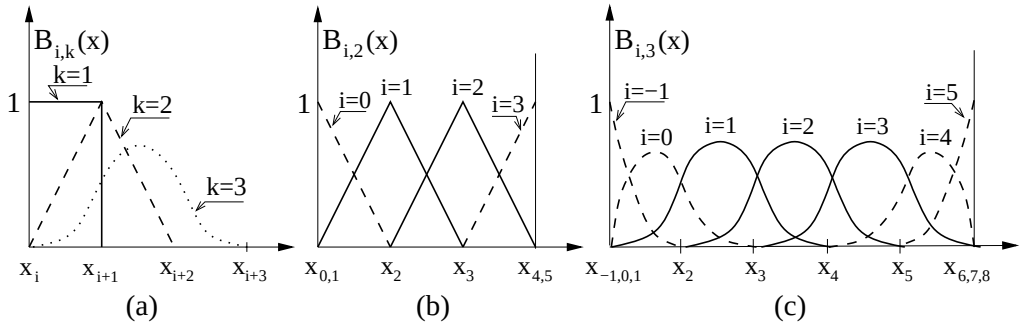


FIGURE 11

L'approximation d'une fonction $f(x)$ par une combinaison linéaire de B-splines d'ordre fixé k s'adapte au cas où $f(x)$ est restreinte à un intervalle fini $[a, b]$ ($x_1 = a$, $x_N = b$). Pour celà on garde la récurrence ci-dessus mais, pour que l'approximation puisse prendre les valeurs

$f(a)$ et $f(b)$ (différentes de zéro en général) en a et b, on ajoute à chaque extrémité de l'intervalle $k-1$ noeuds confondus avec x_1 et x_N. Par exemple pour $k=2$ (figure 11b où $N=4$) on ajoute x_0 confondu avec x_1 et x_{N+1} confondu avec x_N, ce qui introduit les deux fonctions supplémentaires $B_{0,2}(x) = \frac{x_2-x}{x_2-x_1} B_{1,1}(x) \neq 0$ en x_1 et $B_{N-1,2}(x) = \frac{x-x_{N-1}}{x_N-x_{N-1}} B_{N-1,1}(x) \neq 0$ en x_N (les deux fonctions porte $B_{0,1}(x)$ et $B_{N,1}(x)$ également introduites étant nulles car de supports $]x_0, x_1[$ et $]x_N, x_{N+1}[$ "vides"). Pour chaque k on construit ainsi au total $N+k-2$ fonctions $B_{i,k}(x)$ avec i allant de $2-k$ à $N-1$. (Le lecteur vérifiera que $N+k-2 = (N-1)k - (N-2)(k-2)$ est la dimension de l'espace vectoriel des polynômes de degré $k-1$ et dérivables $k-2$ fois en $x_2, x_3 \cdots x_{N-1}$.) Les fonctions $B_{i,k}(x)$ sont alors nulles en dehors de $[a,b]$ et vérifient toujours $\sum_i B_{i,k}(x) = 1$. Pour $i = 2-k, 3-k, \cdots, 1$, elles partent toutes de x_1 et ont pour support $[x_1,x_2], [x_1,x_3] \cdots [x_1,x_{k+1}]$; elles sont de plus en plus régulières en x_1 lorsque i augmente (discontinue pour $i = 2-k$, continue à dérivée première discontinue pour $i = 3-k$...) ; les B-splines pour $i = N-1, N-2, \cdots N-k$ ont un comportement similaire en x_N (figure 11c où $N=6$ et $k=3$).

Un intérêt de la représentation de $f(x)$ par des B-splines d'ordre fixé k

$$f(x) \simeq \sum_{i=2-k}^{N-1} c_i \, B_{i,k}(x)$$

est que pour tout x elle n'implique qu'un petit nombre (k au plus) de $B_{i,k}(x)$. Un autre intérêt est qu'elle permet d'écrire simplement les conditions aux limites susceptibles d'être imposées à f en x_1 et x_N ; par exemple, pour $k=3$, on a $c_{-1} = f(a)$ et $c_{N-1} = f(b)$. Enfin, comme pour les familles de polynômes orthogonaux, il existe des relations remarquables entre les $B_{i,k}$ et leurs dérivées, par exemple :

$$B'_{i,k}(x) = (k-1)\Big(\frac{B_{i,k-1}(x)}{x_{i+k-1}-x_i} - \frac{B_{i+1,k-1}(x)}{x_{i+k}-x_{i+1}}\Big) \, .$$

L'ensemble de ces propriétés fait que la représentation de f par des B-splines est bien adaptée à la résolution d'E.D. avec conditions aux limites.

11.3.2 Interpolation de Lagrange et par "cubic-splines"

▷ Polynômes de Lagrange

Par N points $(x_i, y_i = f(x_i))$ du plan (x,y) il ne passe qu'une seule courbe polynomiale $y = P(x)$ de degré $N-1$. $P(x)$ est donné par la formule explicite de **Lagrange**

$$P(x) = \sum_{i=1}^{N} f(x_i)\, l_i(x) \quad \text{avec } l_i(x) = \prod_{j \neq i} \frac{x-x_j}{x_i-x_j} \, .$$

$P(x)$ peut aussi être représenté sur la base 1, $(x-x_1)$, $(x-x_1)(x-x_2), \cdots$ par

$$P(x) = f(x_1) + \sum_{k=1}^{N} f[x_1, x_2, \cdots, x_k]\,(x-x_1)(x-x_2)\cdots(x-x_k)$$

où les coefficients $f[x_1, x_2, \cdots, x_k]$, appelés **différences divisées**, sont donnés par la récurrence

$$f[x_1, x_2, \cdots, x_k] = \frac{f[x_2, x_3, \cdots, x_k] - f[x_1, x_2, \cdots, x_{k-1}]}{x_k - x_1} \quad \text{avec} \quad f[x_i] = f(x_i) \, .$$

Ces expressions généralisent l'interpolation linéaire pour deux points ($N = 2$) :

$$P_{12}(x) = \frac{(x_2 - x)f(x_1) + (x - x_1)f(x_2)}{x_2 - x_1} = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \ .$$

En pratique $P(x)$ est obtenu en définissant à chaque étape des polynômes qui interpolent ceux de l'étape précédente (méthode de **Neville**).

EXEMPLE ($N = 3$) (figure 12). On part à l'étape 0 des polynômes constants $P_i = f(x_i)$ ($i = 1, 2, 3$). On construit, à l'étape 1 les interpolations linéaires $P_{12}(x)$ (expression ci-dessus) et $P_{23}(x)$, puis à l'étape 2 le polynôme $P_{123}(x) = (x_3 - x_1)^{-1}\big((x_3 - x)P_{12}(x) + (x - x_1)P_{23}(x)\big)$ obtenu à partir de P_{12} et P_{23} comme l'est P_{13} à partir de P_1 et P_3).

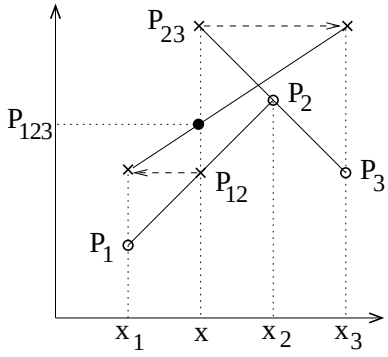


FIGURE 12

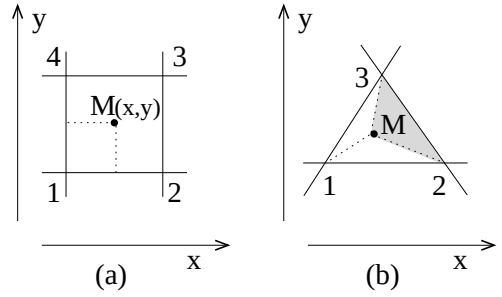


FIGURE 13

Interpolation en dimension 2 (figure 13). Pour une grille rectangulaire elle est donnée par $f(x, y) = \lambda_x\lambda_y f_1 + (1 - \lambda_x)\lambda_y f_2 + \lambda_x(1 - \lambda_y)f_3 + (1 - \lambda_x)(1 - \lambda_y)f_4$ ($\lambda_x = (x_2 - x)(x_2 - x_1)^{-1}$, $\lambda_y = (y_3 - y)(y_3 - y_2)^{-1}$ et $f_i = f(x_i, y_i)$) ; pour une grille triangulaire

$$f(x, y) = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$$

où les λ_i sont les coordonnées barycentriques ($\lambda_1 = \dfrac{\text{aire } M23}{\text{aire } 123} \cdots$, *cf.* section 2.2.1).

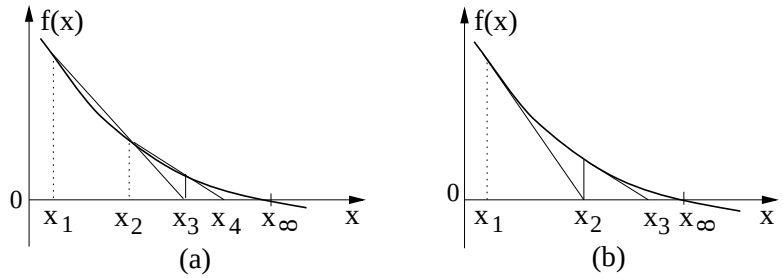


FIGURE 14

Interpolation et recherche de zéros. Supposons que x_1 et x_2 sont proches d'une valeur x_∞ où $f(x)$ s'annule. En posant $P_{12}(x) = 0$ dans l'interpolation linéaire ci-dessus récrite $x = x_1 + \dfrac{P_{12}(x) - f(x_1)}{f(x_2) - f(x_1)}(x_2 - x_1)$ on trouve une valeur approchée $x_3 = x_1 -$

$\dfrac{f(x_1)}{f(x_2)-f(x_1)}(x_2-x_1)$ pour laquelle $f(x)=0$. On itère cette méthode en remplaçant (x_1,x_2) par (x_2,x_3) pour obtenir x_4 *etc.* (méthode de la **sécante** ; figure 14a). Cette méthode remplace celle de **Newton-Raphson** (figure 14b) $x_{i+1}=x_i-\dfrac{f(x_i)}{f'(x_i)}$ lorsque $f'(x_i)$ n'est pas connue avec assez de précision. (Attention, ces méthodes peuvent ne pas converger si on part trop loin de la solution x_∞.)

▷ "Cubic-splines"

Vouloir interpoler entre un très grand nombre de points avec un seul polynôme (nécessairement de degré élevé) n'est pas raisonnable car les oscillations entre points augmentent avec le degré. Il est préférable, notamment pour avoir des graphes "lisses", de raccorder des interpolations polynomiales "locales" en assurant la continuité jusqu'à la dérivée seconde aux points de raccordement.

Construction : deux points x_1 et $x_2=x_1+h$ étant donnés, soient $\lambda_1=h^{-1}(x_2-x)$ et $\lambda_2=h^{-1}(x-x_1)$ les coordonnées barycentriques d'un point intermédiaire x et $f(x_i)=f_i$. La fonction cubique définie sur $[x_1,x_2]$ par

$$S(x)=\lambda_1 f_1+\frac{a_1}{6}(\lambda_1^3-\lambda_1)h^2+\lambda_2 f_2+\frac{a_2}{6}(\lambda_2^3-\lambda_2)h^2$$

vérifie $S(x_i)=f_i$ et $S''(x)=\lambda_1 a_1+\lambda_2 a_2$, d'où $S''(x_i)=a_i$. L'interpolation obtenue par recollement de ces fonctions, définies chacune sur un intervalle $[x_i,x_{i+1}]$, est donc continue ainsi que sa dérivée seconde. Reste à assurer la continuité de la dérivée première en chaque point x_i, ce qui donne un système d'équations pour les a_i. Un petit calcul conduit à la récurrence d'ordre 2 :

$$\frac{x_{i+1}-x_i}{6}\,a_{i+1}+\frac{x_{i+1}-x_{i-1}}{6}\,a_i+\frac{x_i-x_{i_1}}{6}\,a_{i-1}=\frac{f_{i+1}-f_i}{x_{i+1}-x_i}-\frac{f_i-f_{i-1}}{x_i-x_{i-1}}\ .$$

Cette relation a une solution unique si on fixe deux conditions, par exemple $a_1=a_N=0$ (cubic-splines "naturelles") ou en fixant $S'(x_1)$ et $S'(x_N)$. Cette solution est obtenue en pratique par la résolution d'un système linéaire tridiagonal (*cf.* section 11.4.1).

11.3.3 Méthode des moindres carrés

▷ Moindres carrés linéaires

Sous cette forme la plus simple il s'agit d'extraire une forme analytique $f(x)$ d'un ensemble de n données f_i, entachées d'erreur, correspondant aux valeurs x_i d'une variable. Pour ce faire on cherche, comme précédemment, $f(x)$ sous la forme d'une combinaison finie de $p<n$ fonctions de base $\phi_j(x)$ (des polynômes par exemple)

$$f(x)=\sum_{j=1}^{p} c_j\,\phi_j(x)\ ,$$

les coefficients c_j étant ici obtenus en minimisant le carré de l'écart (norme L_2) :

$$e^2=\sum_{i=1}^{n} w_i\big(f_i-\sum_{j=1}^{p} c_j\,\phi_j(x_i)\big)^2 \qquad w_i=\sigma_i^{-2}>0\ .$$

Si on a une estimation des incertitudes sur les f_i, le poids w_i accordé à chacun est d'autant plus grand que l'erreur le concernant est petite. Si les erreurs sur les f_i peuvent

être assimilées à des variables aléatoires gaussiennes indépendantes de variance σ_i^2, e^2 est un χ^2 à n degrés de liberté (*cf.* section 10.2.4). En écrivant

$$\frac{\partial e^2}{\partial c_k} = -2\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\Big(f_i - \sum_{j=1}^{p} c_j\,\phi_j(x_i)\Big)\,\phi_k(x_i) = 0$$

on voit que les c_j vérifient l'équation matricielle :

$$\mathbf{G}^t\mathbf{G}\,\underline{c} = \mathbf{G}^t\underline{b} \qquad (\underline{c}^t = (c_1, \cdots, c_p)) \ .$$

$\mathbf{G}$ est la matrice $n \times p$ d'éléments $G_{ij} = \sigma_i^{-1}\phi_j(x_i)$ et $\underline{b}$ le vecteur à n composantes $b_i = \sigma_i^{-1} f_i$. Malheureusement la matrice $p \times p$ $\mathbf{G}^t\mathbf{G}$ est généralement mal conditionnée (c.a.d. qu'une petite variation dans les données (f_i, x_i) change complètement la solution) et son inversion numérique est alors problématique (*cf.* section 11.4).

EXEMPLE. Pour $\phi_j(x) = x^{j-1}$ (approximation polynomiale), $w_i = 1$ pour tout i et les x_i régulièrement distribués sur $[-1, 1]$ on obtient

$$(\mathbf{G}^t\mathbf{G})_{kj} = \textstyle\sum_{i=1}^{n} x_i^{j+k-2} \simeq n\int_0^1 x^{j+k-2}\,\mathrm{d}x = \frac{n}{j+k-1} \ .$$

$\mathbf{G}^t\mathbf{G}$ est une matrice $p \times p$ de type Hilbert connue pour être mal conditionnée dès que $p \geq 10$.

REMARQUE. Il n'est souvent pas nécessaire de considérer de grandes valeurs de p. Une valeur convenable de p est atteinte si, après avoir diminué rapidement quand p augmente, e^2 minimum ne diminue que lentement au delà de cette valeur.)

▷ Moindres carrés non linéaires

Si la forme analytique $f(x, \underline{a})$ de la fonction qui doit représenter les données est connue par la physique, la méthode des moindres carrés sert à ajuster les valeurs des paramètres $\underline{a} \equiv (a_1, a_2, \cdots a_p)$. On minimise comme précédemment la quantité

$$e^2(\underline{a}) = \sum_{i=1}^{n} w_i\big(f_i - f(x_i, \underline{a})\big)^2 \ ,$$

et on obtient un système de p équations non linéaires pour les a_j

$$\frac{\partial e^2}{\partial a_j} = -2\sum_{i=1}^{n} w_i\big(f_i - f(x_i, \underline{a})\big)\,\frac{\partial f(x_i, \underline{a})}{\partial a_j} = 0 \quad (j = 1, 2, \cdots p)$$

de la forme

$$g_j(a_1, a_2, \cdots a_p) = 0 \quad (\underline{g} : \text{gradient de } e^2) \ .$$

La résolution numérique de ce système s'effectue par une méthode de Newton modifiée. On choisit des valeurs initiales $\underline{a}^{(0)}$ (supposées pas trop éloignées de la solution) et on linéarise le système en écrivant : $g_j(\underline{a}) \simeq g_j(\underline{a}^{(0)}) + \sum_{k=1}^{p}(a_k - a_k^{(0)})\,\dfrac{\partial g_j}{\partial a_k}$.

En notant $\mathbf{J}$ la **matrice hessienne** de e^2, dont les éléments de matrice sont

$$J_{kj} = \frac{\partial^2 e^2}{\partial a_j\,\partial a_k} = \frac{\partial g_j}{\partial a_k} = 2\sum_{i=1}^{n} w_i \frac{\partial f(x_i, \underline{a})}{\partial a_j}\,\frac{\partial f(x_i, \underline{a})}{\partial a_k} - 2\sum_{i=1}^{n} w_i\big(f(x_i, \underline{a}) - f_i\big)\,\frac{\partial^2 f(x_i, \underline{a})}{\partial a_j\,\partial a_k} \ ,$$

le système d'équations linéarisées pour $\underline{a}$ s'écrit (sous forme matricielle) :

$$\underline{g}(\underline{a}^{(0)}) = -\mathbf{J}^{(0)}\,(\underline{a} - \underline{a}^{(0)}) \quad (\mathbf{J}^{(0)} = \mathbf{J} \text{ pour } \underline{a} = \underline{a}^{(0)})\ .$$

Sa solution $\underline{a}^{(1)}$ est la solution approchée à l'ordre 1. Ensuite on itère le processus en remplaçant $\underline{a}^{(0)}$ par $\underline{a}^{(1)}$ dans l'équation matricielle pour obtenir $\underline{a}^{(2)}$ et ainsi de suite :

$$\underline{g}(\underline{a}^{(i)}) = -\mathbf{J}^{(i)}\,(\underline{a}^{(i+1)} - \underline{a}^{(i)})\ .$$

REMARQUE. Près du minimum les différences $\left(f(x_i, \underline{a}) - f_i\right)$ sont petites et non corrélées de telle sorte qu'il est légitime de négliger la deuxième somme dans l'expression de J_{kj}. Dans cette approximation la matrice symétrique $\mathbf{J}$ est le plus souvent définie positive mais peut dans certains cas être presque positive semi-définie ; elle n'est donc pas toujours inversible et pour éviter ce problème on résoud le systéme modifié

$$\underline{a}^{(i+1)} = \underline{a}^{(i)} + \left[\mathbf{J}^{(i)} + \lambda^{(i)}\mathbf{I}\right]^{-1} \underline{g}(\underline{a}^{(i)})$$

où la matrice $\mathbf{J}^{(i)} + \lambda^{(i)}\mathbf{I}$ est définie positive. On vérifie que $e^2(\underline{a}^{(i+1)}) < e^2(\underline{a}^{(i)})$ (sinon on augmente $\lambda^{(i)}$) et on itère le processus en diminuant la valeur de λ à chaque pas pour le faire tendre vers 0 lorsque le degré de convergence est jugé satisfaisant (méthode de **Levenberg-Marquardt**).

11.3.4 Méthode de Bézier

A l'inverse des méthodes d'interpolation et de moindres carrés où le but est de "faire passer au mieux" une fonction par des points fixes donnés, dans la méthode de Bézier, très utilisée en CAO (comception assistée par ordinateur), on déplace les points de contrôle d'une courbe en paramétrique (ou d'une surface) pour obtenir une forme donnée, par exemple celle d'un caractère d'imprimerie (ou d'une carrosserie d'automobile).

▷ Courbes de Bézier

O étant un point fixe quelconque, la courbe de Bézier à $n+1$ points de contrôle P_0, P_1, $\cdots$, P_n est celle décrite, lorsque t parcourt l'intervalle $[0, 1]$, par le point mobile $P(t)$ défini par

$$\overrightarrow{OP(t)} = \sum_{k=0}^{n} C_n^k\, t^k\, (1-t)^{n-k}\, \overrightarrow{OP_k}\ .$$

$P(t)$ est le barycentre des points P_k affectés des poids $B_k^n(t) = C_n^k\, t^k\, (1-t)^{n-k}$ (**polynomes de Berstein** qui vérifient $\sum_{k=0}^{n} B_k^n(t) = 1$).

- Cas de deux points de contrôle (degré 1) : en plaçant 0 en P_0 il vient $\overrightarrow{P_0P(t)} = t\,\overrightarrow{P_0P_1}$; la courbe de Bézier est le segment P_0P_1.

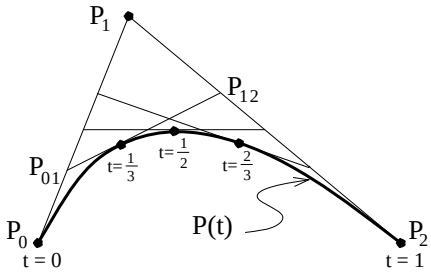


FIGURE 15

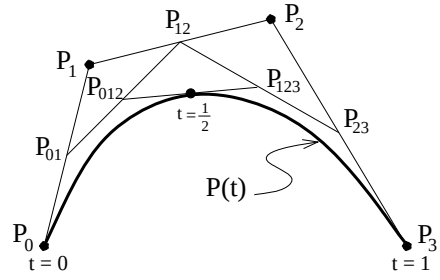


FIGURE 16

- Cas de trois points de contrôle (degré 2) : $\overrightarrow{OP(t)} = (1-t)^2\,\overrightarrow{OP_0} + 2t(1-t)\overrightarrow{OP_1} + t^2\overrightarrow{OP_2}$. Pour tracer la courbe on utilise l'associativité dans la détermination du barycentre : si P_{01} est le barycentre de P_0 (poids $1-t$) et de P_1 (poids t) et si P_{12} est le barycentre de P_1 (poids $1-t$) et P_2 (poids t) alors $P(t)$ est le barycentre de P_{01} (poids $1-t$) et P_{12} (poids t). Lorsque t varie de 0 à 1, P_{01} décrit le segment P_0P_1, P_{12} le segment P_1P_2 et $P(t)$ la courbe de Bézier (figure 15).

Vérification : $\overrightarrow{OP_{01}} = (1-t)\,\overrightarrow{OP_0} + t\,\overrightarrow{OP_1}$ et $\overrightarrow{OP_{12}} = (1-t)\,\overrightarrow{OP_1} + t\,\overrightarrow{OP_2}$ d'où
$$\overrightarrow{OP(t)} = (1-t)\,\overrightarrow{OP_{01}} + t\,\overrightarrow{OP_{12}} = (1-t)^2\,\overrightarrow{OP_0} + 2t(1-t)\overrightarrow{OP_1} + t^2\overrightarrow{OP_2}$$
- Cas de quatre points de contrôle (degré 3) : $\overrightarrow{OP(t)} = (1-t)^3\,\overrightarrow{OP_0} + 3t(1-t)^2\overrightarrow{OP_1} + 3t^2(1-t)\overrightarrow{OP_2} + t^3\overrightarrow{OP_3}$.
Le tracé de la courbe s'obtient comme précédemment, "en prenant le barycentre de deux barycentres de poids $1-t$ et t", mais avec une étape supplémentaire (figure 16).

Dans la pratique on ne considère pas de courbes de Bézier de degré supérieur. Dans le cas d'un grand nombre de points de contrôle on préfére raccorder, en certains de ces points, des courbes de degré deux ou trois. La raison principale est la suivante. Une courbe de Bézier a un caractère global : le déplacement d'un point de contrôle affecte toute la courbe (*cf.* la figure 17 où P_1 est déplacé en P'_1 et P(t) transformée en P'(t)). Sur la figure 18, où la courbe en pointillés est celle associée aux 7 points de contrôle, la courbe en trait plein est constituée de deux (sous)courbes P(t) et P'(t) (chacune de degré 3 construite sur les 4 points de contrôle $P_0P_1P_2P_3$ et $P'_0P'_1P'_2P'_3$) raccordées en $P_3 \equiv P'_0$; sur cette courbe fractionnée on peut déplacer P_1 ou P_0 (et amener ce dernier en P_3 par exemple pour transformer P(t) en un lobe fermé) sans modifier la partie P'(t). En considérant un courbe fractionnée on acquiert donc la possibilité de corriger localement la forme de la courbe. Une courbe fractionnée sera "lisse" (dérivée continue) au point de raccord si celui-ci est situé sur le segment défini par le point qui le précède et celui qui le suit ; si de plus il est au milieu de ce segment, la dérivée seconde est aussi continue. Ceci résulte du fait que pour tout t les vecteurs $\overrightarrow{P_{01}P_{12}}$ (confondu avec $\overrightarrow{P_0P_1}$ pour $t=0$ et $\overrightarrow{P_1P_2}$ pour $t=1$; figure 15) et $\overrightarrow{P_{012}P_{123}}$ (confondu avec $\overrightarrow{P_0P_1}$ pour $t=0$ et $\overrightarrow{P_2P_3}$ pour $t=1$; figure 16) sont proportionnels au vecteur tangent $\frac{\mathrm{d}\overrightarrow{OP(t)}}{\mathrm{d}t}$. (Le lecteur montrera que $\frac{\mathrm{d}\overrightarrow{OP(t)}}{\mathrm{d}t} = 2\overrightarrow{P_{01}P_{12}}$ pour une courbe de degré 2 et que $\frac{\mathrm{d}\overrightarrow{OP(t)}}{\mathrm{d}t} = 3\overrightarrow{P_{012}P_{123}}$ pour le degré 3.)

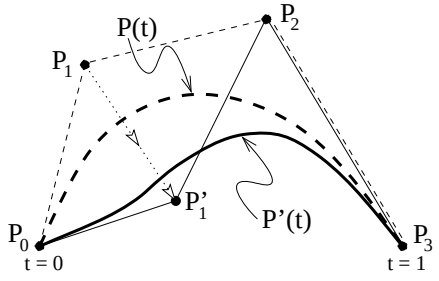


FIGURE 17

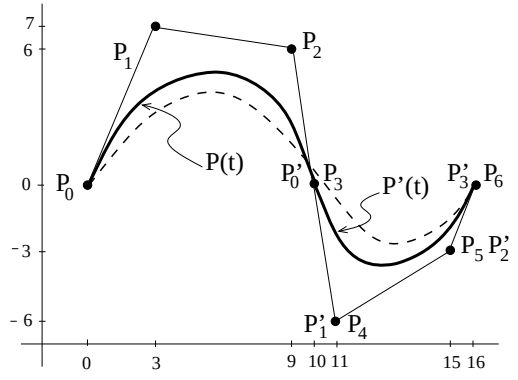


FIGURE 18

REMARQUE. **Courbes B-splines**. Les B-splines $B_{i,k}(t)$ forment une base dans l'espace des polynomes de degré $k-1$ et satisfont $\sum_i B_{i,k}(t) = 1$ comme les polynomes de Berstein. On peut donc, par analogie avec les courbes de Bézier, utiliser la relation

$$\overrightarrow{OP(t)} = \sum_{i=2-k}^{N-1} B_{i,k}(t)\,\overrightarrow{OP_i}$$

pour décrire des courbes en paramétrique : $P(t)$ est le barycentre des $N+k-2$ points de contrôle P_i affectés des poids $B_{i,k}(t)$. Ces courbes B-splines ont un caractère local : les $B_{i,k}$ ayant un support limité un déplacement de P_i ne modifie la courbe qu'au voisinage de ce point. Elles sont, comme les courbes de Bézier, très utilisées en CAO.

▷ Surfaces de Bézier

Pour engendrer une surface, une généralisation naturelle consiste à considérer le point $P(u,v)$ défini par la relation :

$$\overrightarrow{OP(u,v)} = \sum_{k=0}^{n}\sum_{l=0}^{m} C_n^k\,C_m^l\,u^k\,(1-u)^{n-k}\,u^k\,v^l(1-v)^{m-l}\overrightarrow{OP_{kl}}\ ;$$

il décrit la surface de Bézier associée aux $(n+1) \times (m+1)$ points de contrôle P_{kl} lorsque les deux paramètres u et v varient entre 0 et 1. Les deux courbes de Bézier associées aux groupes de points de contôle $\{P_{0l}\}$ et $\{P_{nl}\}$ $(l = 0, 1, \cdots, m)$ limitent la surface en "latitude" tandis que celles associées aux points de contrôle $\{P_{k0}\}$ et $\{P_{km}\}$ $(k = 0, 1, \cdots, n)$ la limitent en "longitude" ("latitude" et "longitude" étant interchangeables).

11.4 RESOLUTION D'EQUATIONS, D'E.D. et D'E.D.P. LINEAIRES

Nous avons vu dans ce chapitre que l'étude numérique de nombreux problèmes (interpolation par des cubic splines, méthode des moindres carrés linéaires ou linéarisés, diverses méthodes aux différences finies appliquées aux E.D. ou E.D.P.) conduit à un système d'équations algébriques linéaires du type $\mathbf{A}\underline{x} = \underline{b}$. Nous allons voir à la section 11.4.2 qu'il en est de même lors de la résolution d'E.D. ou d'E.D.P. par une méthode spectrale. Un tel système linéaire, étudié au chapitre 4, se résoud théoriquement en inversant la matrice $\mathbf{A}$ ou en la diagonalisant, mais le problème n'est pas simple car la dimension de $\mathbf{A}$ peut être grande (par exemple supérieure à 100×100). Heureusement dans de nombreux cas la matrice $\mathbf{A}$ est proche d'une matrice diagonale (souvent tri-diagonale) et des algorithmes d'inversion efficaces peuvent alors lui être appliqués. Cependant il arrive que des lignes (ou des colonnes) de $\mathbf{A}$ soient presque proportionnelles (linéairement dépendantes) ; le problème est alors dit mal conditionné et sa résolution relève de la décomposition S.V.D. (singular value decomposition).

11.4.1 Equations linéaires régulières et singulières

▷ Système régulier ; "triangularisation"

Soit à résoudre le système de n équations (indépendantes) à n inconnues

$$\sum_{j=1}^{n} a_{ij}\, x_j = b_i \quad (i = 1, \cdots, n) \quad \text{ou} \quad \mathbf{A}\underline{x} = \underline{b}\ .$$

Puisque $\underline{x} = \mathbf{A}^{-1}\underline{b}$ on pourrait penser que le problème numérique consiste à déterminer $\mathbf{A}^{-1}$. Cependant celà demande plus d'opérations (donc plus de temps de calcul et une augmentation des erreurs d'arrondi) que les méthodes par "triangularisation".

Dans la **méthode du pivot de Gauss**, les inconnues sont numérotées de telle sorte que dans les n équations

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1\ , \cdots\cdots\ ,\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n$$

le pivot $a_{11} \neq 0$ soit le plus grand possible (en valeur absolue). On élimine x_1 des équations $2, 3, \cdots, n$ en le remplaçant par son expression tirée de la première pour obtenir le nouveau système

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \quad \text{et} \quad a_{i2}^{(1)}x_2 + a_{i3}^{(1)}x_3 + \cdots + a_{in}^{(1)}x_n = b_i^{(1)} \quad (i = 2, \cdots, n)\ .$$

On itère ensuite le processus : renumérotation éventuelle des $n-1$ variables restantes ("pivotage") pour que le pivot $a_{22}^{(1)} \neq 0$ soit le plus grand possible (en valeur absolue),

élimination de x_2 des équations $3, 4, \cdots, n$ en le remplaçant par son expression tirée de la seconde *etc.* On obtient finalement le système "triangulaire"

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2n}^{(1)}x_n &= b_2^{(1)} \\ a_{33}^{(2)}x_3 + \cdots + a_{3n}^{(2)}x_n &= b_2^{(2)} \\ \ddots \qquad\qquad & \vdots \\ a_{nn}^{(n-1)}x_n &= b_n^{(n-1)} \end{aligned} ,$$

qui permet d'obtenir successivement x_n, $x_{n-1} \cdots$.

La **décomposition LU** d'une matrice $\mathbf{A}$ consiste à trouver une matrice $\mathbf{L}$ (**triangulaire inférieure**, d'où L pour "lower") et une matrice $\mathbf{U}$ **triangulaire supérieure** (U pour "upper") telles que :

$$\mathbf{A} = \mathbf{LU} \ .$$

Le lecteur se convaincra, en considérant par exemple la décomposition d'une matrice 3×3

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & 1 \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix} ,$$

qu'en posant $L_{ii} = 1$ on obtient successivement et par alternance les colonnes de $\mathbf{U}$ et $\mathbf{L}$: $U_{11} = a_{11}$, $L_{i1} = U_{11}^{-1}a_{i1}$, U_{12} et U_{22}, $L_{i2} \cdots$ (relations à la base de l'algorithme de Crout). La solution de $\mathbf{A}\underline{x} = \mathbf{LU}\underline{x} = \underline{b}$ s'obtient alors en résolvant les systèmes triangulaires $\mathbf{L}\underline{c} = \underline{b}$ puis $\mathbf{U}\underline{x} = \underline{c}$. Du point de vue efficacité (temps de calcul et précision) elle est comparable à la méthode du pivot de Gauss mais elle permet en plus, connaissant $\mathbf{L}$ et $\mathbf{U}$, d'obtenir de façon "économique" le déterminant de $\mathbf{A}$ (produit des U_{ii}).

REMARQUES. 1) Si $\mathbf{A} = \mathbf{T}$ est **tridiagonale**, $\mathbf{L}$ et $\mathbf{U}$ sont très simples (bidiagonales). 2) On peut aussi introduire des matrices $\mathbf{L}$ et $\mathbf{U}$ telles que $\mathbf{LA} = \mathbf{U}$ en prenant pour $\mathbf{L}$ un produit de transformations de Householder (voir ci-dessous). 3) Si $\mathbf{A}$ est symétrique définie positive, on utilise aussi la décomposition de Cholevsky $\mathbf{A} = \mathbf{LL}^t$ ($\mathbf{L}$ triangulaire inférieure).

▷ Problème mal conditionné ; méthode T.S.V.D.

EXEMPLE. Considérons la minimisation d'une forme quadratique positive $F(x_1, x_2) = \frac{1}{2}(\alpha x_1^2 + 2\beta x_1 x_2 + \gamma x_2^2) - b_1 x_1 - b_2 x_2$ (avec $\alpha > 0$, $\gamma > 0$ et $\alpha\gamma - \beta^2 = \epsilon > 0$). Elle conduit au système linéaire

$$\begin{aligned} \alpha x_1 + \beta x_2 &= b_1 \\ \beta x_1 + \gamma x_2 &= b_2 \end{aligned} \qquad \text{ou} \qquad \mathbf{A}\underline{x} = \underline{b} \ ,$$

et peut s'interpréter comme la détermination des coordonnées du point M d'intersection de deux droites D_1 et D_2 dans le plan (x_1, x_2). Si la forme est presque dégénérée ($\epsilon \simeq 0$), les lignes de $\mathbf{A}$ sont quasi proportionnelles et les droites quasi parallèles ; ceci entraine que la position du minimum M est mal définie car très sensible à des petites variations des paramètres de F. Dit autrement, F varie très peu dans la direction du vecteur $\begin{pmatrix} \beta \\ -\alpha \end{pmatrix}$ associé à la valeur propre proche de 0 de $\mathbf{A}$ et la minimisation dans

cette direction est a priori non pertinente (si on n'a pas d'autres informations sur le problème physique qui a conduit à F). Cette direction est celle du grand axe des ellipses très allongées d'équations $F(x_1, x_2) = C^{\text{ste}}$ qui dessinent les lignes de niveau d'une vallée plutôt que celles d'un puits (figure 19). La **méthode T.S.V.D.** consiste à restreindre la minimisation à l'espace orthogonal (direction du petit axe des ellipses) qui est engendré par le vecteur $\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ associé à la valeur propre finie de $\mathbf{A}$ (point M_R).

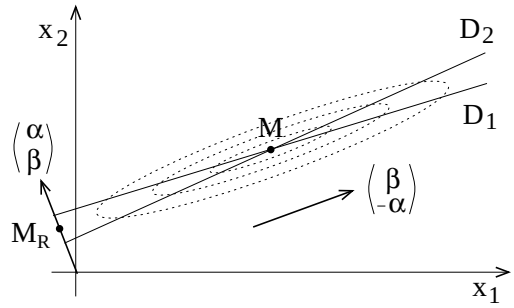


FIGURE 19

Un système linéaire est dit mal conditionné si le rapport $K = \left|\frac{\lambda_{\max}}{\lambda_{\min}}\right|$ des valeurs propres extrêmes de $\mathbf{A}$ est très supérieur à 1 (propriété qui reste vraie, comme la quasi-dépendance linéaire de lignes ou de colonnes, si on multiplie $\mathbf{A}$ par une constante). De la relation

$$\frac{\| \delta \underline{x} \|}{\| \underline{x} \|} = \frac{\| \mathbf{A}^{-1} \delta \underline{b} \|}{\| \mathbf{A}^{-1} \underline{b} \|} \leq K \frac{\| \delta \underline{b} \|}{\| \underline{b} \|}$$

(et d'une relation analogue si $\mathbf{A}$ fluctue) on déduit que, pour $K \gg 1$, $\underline{x}$ est très sensible aux erreurs expérimentales sur $\underline{b}$ et $\mathbf{A}$, et aux erreurs numériques en cours de calcul. ($\mathbf{A}$ est dite "numériquement singulière" si $K \dfrac{\delta x}{x} \geq 1$, où $\dfrac{\delta x}{x}$ est l'erreur relative d'arrondi.) En grande dimension il est très important d'identifier ce type de problème avant de se lancer dans la résolution numérique d'un système linéaire.

La **décomposition S.V.D.** (Singular Value Decomposition) d'une matrice $n \times n$ réelle $\mathbf{A}$ s'écrit

$$\mathbf{A} = \sum_{i=1}^{n} \sigma_i \underline{u_i}\, \underline{v_i}^t$$

où les $\underline{u_i}$ et les $\underline{v_i}$ sont les vecteurs propres orthonormés des matrices (symétriques, définie-positives) $\mathbf{A}\mathbf{A}^t$ et $\mathbf{A}^t\mathbf{A}$ associés aux valeurs propres $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 > 0$. Cette décomposition est unique si tous les σ_i sont distincts. On voit sur l'expression de l'inverse de $\mathbf{A}$

$$\mathbf{A}^{-1} = \sum_{i=1}^{n} \sigma_i^{-1} \underline{v_i}\, \underline{u_i}^t$$

($\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ car $\underline{u_i}^t\, \underline{u_j} = \underline{v_i}^t\, \underline{v_j} = \delta_{ij}$ et $\sum_i \underline{u_i}\, \underline{u_i}^t = \sum_i \underline{v_i}\, \underline{v_i}^t = \mathbf{I}$) qu'elle permet de détecter si un système $\mathbf{A}\underline{x} = \underline{b}$ est mal conditionné : il correspond à $\sigma_{p+1}, \cdots, \sigma_n$

proches de 0. La solution $\underline{x} = \mathbf{A}^{-1}\underline{b}$ contient alors une **partie singulière** $\underline{x_S} = \sum_s \sigma_s^{-1} \underline{v_s}\,\underline{u_s}^t \underline{b}$ $(s = p+1, \cdots, n)$, et une **partie régulière**

$$\underline{x_R} = \sum_r \sigma_r^{-1} \underline{v_r}\,\underline{u_r}^t \underline{b} \quad (r = 1, \cdots, p)$$

qui est la **solution T.S.V.D.** du problème (T pour "truncated" puisque les valeurs σ_i avec $i > p$ n'apparaissent pas). Si on pose $\mathbf{A}_T = \sum_r \sigma_r \underline{u_r}\,\underline{v_r}^t$ on vérifie que $\underline{x_R}$ est le vecteur de norme minimum qui minimise $\| \mathbf{A}_T\underline{x} - \underline{b} \|^2$ (point M_R de la figure 19).

DÉMONSTRATION. $\mathbf{A}_T\underline{x_R} = \sum_r \underline{u_r}\,\underline{u_r}^t\underline{b}$ entraîne $\| \mathbf{A}_T(\underline{x_R} + \underline{y}) - \underline{b} \|^2 = \| -\sum_s \underline{u_s}\,\underline{u_s}^t\underline{b} + \mathbf{A}_T\underline{y} \|^2 = \| \mathbf{A}_T\underline{x_R} - \underline{b} \|^2 + \| \mathbf{A}_T\underline{y} \|^2$ (car $\mathbf{A}_T\underline{y}$ est une combinaison de vecteurs $\underline{u_r}$) $\geq \| \mathbf{A}_T\underline{x_R} - \underline{b} \|^2$. L'égalité a lieu si $\underline{y} \in \mathrm{Ker}\mathbf{A}_T$ est une combinaison des vecteurs $\underline{u_s}$ et alors $\| \underline{x_R} + \underline{y} \|^2 = \| \underline{x_R} \|^2 + \| \underline{y} \|^2 \geq \| \underline{x_R} \|^2$. $\underline{x_R}$ est donc, parmi les vecteurs qui minimisent $\| \mathbf{A}_T\underline{x} - \underline{b} \|^2$, celui qui a la norme minimum.

REMARQUES. 1) **Problèmes surdéterminés**. Si $\mathbf{A}$ est une matrice $m \times n$ avec $m > n$ le système $\mathbf{A}\underline{x} = \underline{b}$ contient plus d'équations (plus de données b_i) que d'inconnues et n'a pas de solution. La décomposition S.V.D. $\mathbf{A} = \sum_i \sigma_i \underline{u_i}\,\underline{v_i}^t$ existe encore mais les $\underline{u_i}$ sont les vecteurs propres orthonormés (à m composantes) de $\mathbf{A}\mathbf{A}^t$ associés aux n valeurs propres non nulles $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 > 0$. La solution T.S.V.D. $\underline{x_R}$ est donnée par la même expression que ci-dessus et minimise $\| \mathbf{A}_T\underline{x} - \underline{b} \|^2$. Si $\mathbf{A}$ n'est pas singulier, $\mathbf{A}_T = \mathbf{A}$ et $\underline{x_R} = (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\underline{b}$.

2) Il existe des stratégies variées basées sur la décomposition S.V.D. et permettant de régulariser un problème d'inversion mal conditionné. On peut minimiser $\| \mathbf{A}_T\underline{x} - \underline{b} \|^2$ en profitant de l'indétermination sur $\underline{x}$ pour imposer une condition plus adaptée au problème physique que "$\| \underline{x} \|^2$ minimum". On peut aussi remplacer $\mathbf{A}^{-1}$ par $\sum_i p(\lambda, \sigma_i)\,(\sigma_i^{-1}\underline{v_i}\,\underline{u_i}^t)$ où p est une fonction poids qui ne coupe pas brutalement les σ_i proches de 0 ; c'est ce que fait en particulier la **méthode de Tikhonov** qui consiste à minimiser $\| \mathbf{A}_T\underline{x} - \underline{b} \|^2 + \lambda^2 \| \underline{x} \|^2$, et donc à remplacer $\mathbf{A}^t\mathbf{A}$ par $\mathbf{A}^t\mathbf{A} + \lambda^2\mathbf{I}$ ou σ_i^2 par $\sigma_i^2 + \lambda^2$.

La décomposition S.V.D. requiert la détermination des valeurs et vecteurs propres de matrices ($\mathbf{A}^t\mathbf{A}$ et $\mathbf{A}\mathbf{A}^t$ qui sont symétriques positives) ; c'est l'objet de l'algorithme QR.

▷ Diagonalisation ; algorithme QR

En théorie les valeurs propres λ_i d'une matrice $\mathbf{A}$ sont obtenues en déterminant les racines du polynôme caractéristique $P(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$ (et les vecteurs propres $\underline{v_i}$ en résolvant les équations $\mathbf{A}\underline{v_i} = \lambda_i\underline{v_i}$). Cependant, du point de vue numérique, il existe des méthodes beaucoup plus efficaces pour trouver les λ_i. Nous décrivons dans la suite la plus couramment utilisée en nous limitant au cas des matrices réelles symétriques.

Dans un premier temps on transforme $\mathbf{A}$ en une matrice tridiagonale $\mathbf{T}$ (symétrique) par des transformations successives de **Householder**. Celles-ci $\mathbf{A} \to \mathbf{A}_1 = \mathbf{P}_1\mathbf{A}\mathbf{P}_1 \to \mathbf{A}_2 = \mathbf{P}_2\mathbf{A}_1\mathbf{P}_2 \cdots$, avec $\mathbf{P}_i = \mathbf{P}_i^t = \mathbf{P_i}^{-1}$, font apparaître les zéros de $\mathbf{T}$ successivement dans les premières colonne et ligne, secondes colonne et ligne... Par exemple, après deux transformations :

$$\mathbf{A} = \begin{pmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix} \longrightarrow \mathbf{A}_2 = \begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & * & * & * & * \end{pmatrix}.$$

La matrice $\mathbf{P}_r$ qui fait passer de $\mathbf{A}_{r-1}$ à $\mathbf{A}_r$ est de la forme $\mathbf{P}_r = \mathbf{I} - 2\underline{u_r}\,\underline{u_r}^t$, où $\underline{u_r} = \frac{\underline{w_r}}{||\underline{w_r}||}$ est un vecteur unitaire dont les r premières composantes sont nulles ; par exemple $\mathbf{P}_3$ doit vérifier :

$$\mathbf{A}_3 = \mathbf{P}_3\mathbf{A}_2\mathbf{P}_3 = \begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ 0 & * & * & * & 0 & 0 \\ 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * & * \end{pmatrix} \quad \text{où} \quad \mathbf{P}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & P_{43} & 1 & 0 & 0 \\ 0 & 0 & P_{53} & 0 & 1 & 0 \\ 0 & 0 & P_{63} & 0 & 0 & 1 \end{pmatrix} .$$

On montre que les $n-r$ composantes non nulles de $\underline{w}_r$ qui conviennent sont : $w_{r,i} = a_{r\,r+1} + \text{signe}\,(a_{r\,r+1})\,(\sum_{i=r+1}^{n} a_{ri}^2)^{\frac{1}{2}}$ si $i = r+1$ et $w_{r,i} = a_{ri}$ pour $i = r+2, \cdots, n$ (a_{ij} éléments de matrices de $\mathbf{A}$) ; les éléments de matrice P_{ir} $(i > r)$ de $\mathbf{P}_r$ s'en déduisent. En opérant ainsi $n-2$ fois on transforme $\mathbf{A}$ en une matrice tridiagonale $\mathbf{T}$.

Dans un deuxième temps les éléments de part et d'autre de la diagonale de $\mathbf{T}$ sont rendus de plus en plus petits par une succession de transformations qui constituent l'algorithme QR. Il consiste à factoriser à chaque étape "s" la matrice $\mathbf{T}_s$ en une matrice orthogonale $\mathbf{Q}_s$ et une matrice triangulaire supérieure $\mathbf{R}_s$ (R pour triangle "right") puis à renverser l'ordre du produit pour obtenir $\mathbf{T}_{s+1}$:

$$\mathbf{T}_s = \mathbf{Q}_s\mathbf{R}_s \quad \text{et} \quad \mathbf{T}_{s+1} = \mathbf{R}_s\mathbf{Q}_s = \mathbf{Q}_s^t\mathbf{T}_s\mathbf{Q}_s \qquad (\mathbf{T}_1 = \mathbf{T}) .$$

On montre que si $\mathbf{T}$ a des valeurs propres distinctes, $\mathbf{T}_s$ tend vers une matrice diagonale ayant pour éléments les valeurs propres de $\mathbf{A}$ (rangées par ordre croissant de module), les éléments non diagonaux $(\mathbf{T}_s)_{i\,i+1} = (\mathbf{T}_s)_{i+1\,i}$ tendant vers 0 comme $\left(\frac{\lambda_i}{\lambda_{i+1}}\right)^s$.

11.4.2 E.D. et E.D.P. linéaires ; méthodes spectrales ; éléments finis

▷ Principe des méthodes spectrales

La résolution numérique par une méthode spectrale d'une E.D., par exemple

$$y''(x) = F(y, y', x) ,$$

consiste à approcher y par un développement fini sur une base de fonctions connues (fonctions sinusoïdales, polynômes, B-splines...) $y(x) \simeq \sum_{p=1}^{N} c_p\,\phi_p(x)$ et à obtenir les coefficients c_p en portant ce développement dans les deux membres de l'E.D. (avec $y' = \sum_p c_p\,\phi_p'$ et $y'' = \sum_p c_p\,\phi_p''$).

Celle-ci ne peut pas être satisfaite exactement mais peut l'être de façon approchée
- soit par une **méthode de collocation** qui consiste à restreindre l'égalité $y'' = F$ à un ensemble discret de points $y''(x_i) = F(y(x_i), y'(x_i), x_i)$. (Comme dans l'intégration de Gauss, il peut y avoir des choix des x_i plus judicieux que d'autres, selon la base $\{\phi_p\}$.)
- soit par la **méthode de Galerkin** (ou des **éléments finis**) qui consiste à projeter l'E.D. sur la base à l'aide d'un produit scalaire :

$$< \phi_q, \sum_{p=1}^{N} c_p\,\phi_p'' > = < \phi_q, F > \quad (q = 1, \cdots, N) .$$

Notons que certains coefficients c_p peuvent ne plus apparaître comme des inconnues si ils ont servi à satisfaire des C.L. (ou des conditions de normalisation pour des solutions ondes stationnaires). Inversement d'autres inconnues que les c_p peuvent être présentes dans l'E.D., par exemple ω pour les ondes stationnaires (*cf.* section 11.2.2). Dans tous

les cas on est conduit à résoudre un système d'équations linéaires comportant un grand nombre d'inconnues. Ceci est évident si F est linéaire (et si ω n'apparaît pas comme inconnue). Dans le cas général notons $\underline{a}$ les inconnues (qui ne se limitent pas forcément aux c_p) et $g_q(\underline{a}) = 0$ les équations à satisfaire. Leur résolution par linéarisation est analogue à celle vue à la section 11.3.3.

REMARQUE. La résolution d'une E.D. par des différences finies, où on remplace par exemple y'' par $h^{-2}(y_{i+1} - 2y_i + y_{i-1})$ et y' par $(2h)^{-1}(y_{i+1} - y_{i-1})$ avec $y_i = y(x_i)$, conduit aussi à de tels systèmes linéaires pour les y_i.

▷ Eléments finis à une dimension

EXEMPLE. Considérons, sur un intervalle $[a, b]$, l'E.D.

$$-y''(x) + f(x)\, y(x) = g(x)$$

et projetons le membre de gauche sur un E.V. de fonctions test ϕ suffisamment régulières (ϕ continue sur $[a, b]$) engendré par des fonctions $\{\phi_q\}$. Une intégration par parties donne :

$$\int_a^b \phi \left(-y'' + fy\right) \mathrm{d}x = \int_a^b (\phi' y' + f\phi y)\, \mathrm{d}x + \phi(a)\, y'(a) - \phi(b)\, y'(b) \ .$$

Le terme $\phi(a)\, y'(a)$ disparait si $y'(a) = 0$ (C.L. de Neuman en a pour $y(x)$, auquel cas $\phi(a)$ est arbitraire) ou si $\phi(a) = 0$ (condition imposée à $\phi(x)$ lorsque $y(x)$ obéit à la C.L. de Dirichlet $y(a) = 0$) ; il en est de même pour le terme $\phi(b)\, y'(b)$. Avec ces conditions et après projection, l'E.D. conduit à :

$$\int_a^b (\phi' y' + f\phi y)\, \mathrm{d}x = \int_a^b \phi g \ \mathrm{d}x \ .$$

Si y est cherché sous la forme $y \simeq \sum_{p=1}^{N} c_p\, \phi_p$, la résolution (approchée) de l'E.D. conduit au système linéaire :

$$\sum_p A_{qp}\, c_p = g_q \quad \text{où} \quad g_q = \int_a^b \phi_q\, g\, \mathrm{d}x \quad \text{et} \quad A_{qp} = \int_a^b (\phi'_q\, \phi'_p + f\, \phi_q\, \phi_p)\, \mathrm{d}x \ .$$

La matrice $\mathbf{A}$ d'éléments A_{qp} a comme propriétés intéressantes d'être symétrique, définie positive si $f(x) > 0$, et proche d'une matrice diagonale si les fonctions ϕ_q et ϕ_p ne se recouvrent plus dès que $|q - p|$ est assez grand. Par exemple si $[a, b]$ est divisé en $N + 1$ intervalles égaux ($x_0 = a$ et $x_{N+1} = b$) et si $y(a) = y(b) = 0$, le choix le plus simple pour ϕ_p est la fonction triangle de hauteur 1 et de base $[x_p, x_{p+2}]$ avec $p = 1, \cdots, N$ (approximation de y par des fonctions linéaires par morceaux). On a alors $A_{p+1\,p} = A_{pp} = A_{p\,p+1} \neq 0$ et $A_{qp} = 0$ sinon (matrice tridiagonale).

Remarque : le caractère symétrique de $\mathbf{A}$ est lié mathématiquement à l'absence de terme (linéaire) en y' dans l'E.D. et physiquement au fait que l'E.D. correspond à l'extrémalisation d'une "énergie" $\int_a^b \left(\frac{1}{2}y'^2 + \frac{1}{2}fy^2 - gy\right) \mathrm{d}x$ dont $\sum_{qp} \frac{1}{2} A_{qp} c_q c_p - \sum_q c_q g_q$ est une forme discrète approchée.

▷ Eléments finis à deux dimensions

EXEMPLE. Considérons, dans un domaine $\mathcal{D}$, l'**E.D.P. de Poisson**

$$-\Delta f(x,y) = g(x,y)$$

où g est un terme de source et où f satisfait sur la frontière γ de $\mathcal{D}$ (figure 20a) soit des C.L. de Dirichlet ($f = C^{\text{ste}}$ sur la partie γ_1 de γ) soit des C.L. de Neuman ($\overrightarrow{\nabla} f \cdot \hat{n} = 0$ sur la partie $\gamma_2 = \gamma - \gamma_1$). A l'aide de la formule de Green on obtient, pour une fonction test ϕ suffisamment régulière

$$\iint_{\mathcal{D}} -(\Delta f)\,\phi \,\mathrm{d}x\,\mathrm{d}y = \iint_{\mathcal{D}} \overrightarrow{\nabla} f \cdot \overrightarrow{\nabla}\phi \,\mathrm{d}x\,\mathrm{d}y - \oint_{\gamma} \phi\, \overrightarrow{\nabla} f \cdot \hat{n}\,\mathrm{d}l \ ,$$

et, comme à une dimension, le terme de bord disparaît si on impose $\phi = 0$ sur γ_1.

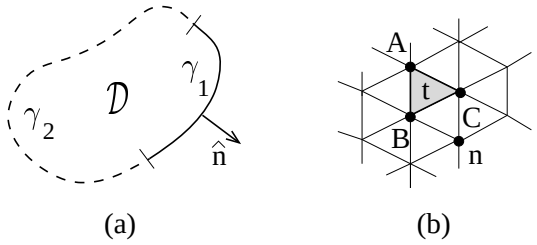


FIGURE 20

Pour appliquer la méthode des éléments finis on découpe $\mathcal{D}$ en éléments simples. Par exemple dans une triangulation les éléments sont des triangles t (disjoints ou ayant en commun soit un sommet soit un coté entier ; figure 20b). Soient n les nœuds "libres" numérotés du réseau, sommets de triangles sur lesquels la valeur de f n'est pas fixée par les C.L.. Dans le cas d'une approximation de f par des fonctions linéaires par morceaux, la valeur $f(x,y)$ de f en un point M intérieur d'un triangle $t = ABC$ direct est approximée par

$$f(x,y) \simeq f(A)\lambda_{tA}(x,y) + f(B)\lambda_{tB}(x,y) + f(C)\lambda_{tC}(x,y)$$

où les λ_{tn} sont les coordonnées barycentriques de M relatives aux nœuds $n = A, B, C$. On a alors :

$$\overrightarrow{\nabla} f = \Delta_t^{-1}\,\hat{z} \wedge \left(f(A)\,\overrightarrow{BC} + f(B)\,\overrightarrow{CA} + f(C)\,\overrightarrow{AB}\right) \quad \text{avec} \quad \Delta_t = 2 \text{ aire}\,(ABC) \ .$$

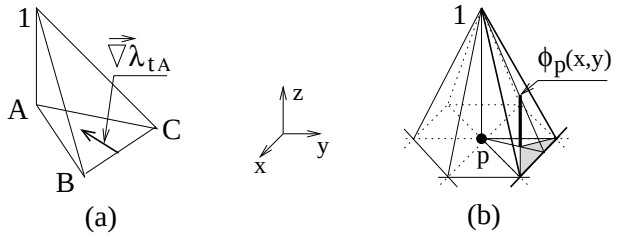


FIGURE 21

(On obtient cette expression en observant que $\overrightarrow{\nabla}\lambda_{tA}$ est constant car λ_{tA} est linéaire, que $\overrightarrow{\nabla}\lambda_{tA} \perp \overrightarrow{BC}$ car λ_{tA} est nulle sur BC et que $\overrightarrow{\nabla}\lambda_{tA} \cdot \overrightarrow{BA} = 1$; λ_{tB} et λ_{tC} satisfont des relations équivalentes ; figure 21a.) L'espace des fonctions ϕ sur lequel on projette l'E.D.P. est engendré par des fonctions ϕ_p elles aussi linéaires par morceaux, égales à 1 au nœud p et nulles aux autres nœuds ; elles se confondent avec λ_{tp} sur les triangles t de sommet libre commun p (figure 21b). Les inconnues $f(n) \equiv f_n$, valeurs

de f aux nœuds libres, sont alors solutions du système linéaire :

$$\sum_n A_{pn} f_n = g_p \quad \text{avec} \quad \sum_n A_{pn} f_n = \iint_{\mathcal{D}} \overrightarrow{\nabla}\phi_p \cdot \overrightarrow{\nabla} f \,\mathrm{d}x\,\mathrm{d}y \quad \text{et} \quad g_p = \iint_{\mathcal{D}} \phi_p \, g \,\mathrm{d}x\,\mathrm{d}y \ .$$

Remarques. 1) $A_{pn} = \sum_t \Delta_t \overrightarrow{\nabla}\phi_p \cdot \overrightarrow{\nabla}\phi_n$ où la somme porte sur les triangles ayant pour sommets les nœuds libres n et p ($A_{pn} = 0$ si les sommets ne sont pas contigus) ; $g_p = \sum_t \frac{1}{6}\Delta_t g_t$ où la somme porte sur les triangles de sommet p et où g est assimilé à une constante g_t sur chaque triangle (l'intégrale de ϕ_p sur le triangle t valant $\frac{1}{6}\Delta_t$).

2) Le système linéaire s'obtient aussi par minimisation de $\iint_{\mathcal{D}} \big(\frac{1}{2}\big(\overrightarrow{\nabla} f\big)^2 - fg\big)\,\mathrm{d}x\,\mathrm{d}y$.

▷ Recherche de valeurs propres par collocation

EXEMPLE. Considérons, dans un domaine $\mathcal{D}$ de frontière γ, l'E.D.P.

$$-\Delta f(x,y) = E\, f(x,y)$$

avec des C.L. de Dirichlet $f = 0$ sur γ. On peut choisir pour les fonctions ϕ_p des ondes planes $\sin(k_x x + k_y y)$ et $\cos(k_x x + k_y y)$ avec $k_x^2 + k_y^2 = k^2 = E$ fixé et des directions de $\overrightarrow{k}$ isotropes. Puisque ces fonctions satisfont l'E.D.P., on détermine leurs coefficients simplement en exprimant que $f = \sum_{p=1}^{N} c_p\, \phi_p = 0$ en M points de γ, et $f = 1$ en un point intérieur fixé. En prenant des points supplémentaires P_i sur γ (environ trois fois plus), on calcule la quantité $I(k) = \sum_i |f(P_i)|^2$ qui devrait être nulle si f était une fonction propre. Lorsque k (donc E) varie, $I(k)$ présente des minima pour les valeurs propres cherchées. En pratique, pour que $I(k)$ varie "régulièrement", il faut non seulement considérer un grand nombre d'ondes planes ($N \sim 100$), mais aussi choisir $N > M + 1$ ce qui implique de résoudre un système linéaire sous déterminé pour les c_p, et donc d'utiliser une méthode S.V.D. (*cf.* section 11.4.1).

11.5 RECHERCHE DE MINIMA

11.5.1 Méthodes utilisant le gradient

▷ Méthodes de la steepest descent et du gradient conjugué

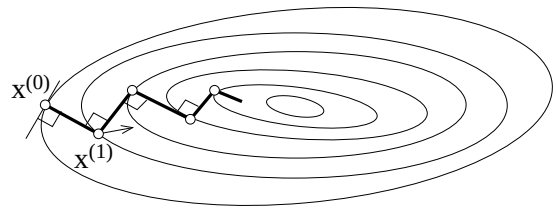


FIGURE 22

La méthode de la **steepest descent** consiste, à l'étape k à partir de $\underline{x}^{(k)}$, à chercher le minimum de $f(\underline{x})$ dans la direction de $-\underline{\partial} f(\underline{x}^{(k)})$ qui est celle dans laquelle f décroit le plus vite. Elle revient donc à déterminer la valeur $\alpha_{\mathrm{m}}^{(k)}$ de $\alpha^{(k)}$ qui minimise $f\big(\underline{x}^{(k)} - \alpha^{(k)} \underline{\partial} f(\underline{x}^{(k)})\big)$, puis à itérer le processus à partir de $\underline{x}^{(k+1)} = \underline{x}^{(k)} - \alpha_{\mathrm{m}}^{(k)} \underline{\partial} f(\underline{x}^{(k)})$. Cette méthode n'est pas très efficace car si il existe une vallée les plus grandes pentes locales ont tendance à y faire descendre le chemin avant d'atteindre le minimum ; par ailleurs le chemin de steepest descent décrit une ligne brisée qui tourne à angle droit en chaque point $\underline{x}^{(k)}$ ce qui n'est pas la meilleure orientation à prendre pour "viser" le minimum (figure 22).

Ces défauts sont corrigés dans la méthode du **gradient conjugué**. Le premier pas de $\underline{x}^{(0)}$ à $\underline{x}^{(1)}$ est toujours effectué dans la direction de plus grande pente. Mais au lieu de tourner à angle droit aux points $\underline{x}^{(k)}$, la direction suivie de $\underline{x}^{(k)}$ à $\underline{x}^{(k+1)}$ $(k \geq 1)$ est donnée par :

$$\underline{s}^{(k)} = -\underline{\partial} f(\underline{x}^{(k)}) + \frac{\underline{\partial} f(\underline{x}^{(k)})^t \underline{\partial} f(\underline{x}^{(k)})}{\underline{\partial} f(\underline{x}^{(k-1)})^t \underline{\partial} f(\underline{x}^{(k-1)})} \underline{s}^{(k-1)} \ .$$

Justification (figure 22). Si $f(\underline{x}) = \frac{1}{2}\underline{x}^t \mathbf{A} \underline{x}$ est une forme quadratique, la direction de départ en $\underline{x}^{(0)}$ est $\underline{s}^{(0)} = -\mathbf{A}\underline{x}^{(0)}$. Le point $\underline{x}^{(1)} = \underline{x}^{(0)} + \alpha^{(0)}\underline{s}^{(0)}$ est déterminé par la condition $\underline{s}^{(0)t}\mathbf{A}\underline{x}^{(1)} = 0$ qui exprime que $\underline{s}^{(0)}$ est orthogonal au gradient de f en $\underline{x}^{(1)}$. La direction idéale de minimisation en $\underline{x}^{(1)}$ n'est pas $-\mathbf{A}\underline{x}^{(1)}$, mais $-\underline{x}^{(1)}$ (appelée **direction conjuguée** de $\underline{s}^{(0)}$ par rapport à $\mathbf{A}$; flèche sur la figure). On peut l'obtenir à deux dimensions (ou s'en rapprocher si $n > 2$) en posant $\underline{s}^{(1)} = -\mathbf{A}\underline{x}^{(1)} + \mu^{(0)}\underline{s}^{(0)}$ et en imposant $\underline{s}^{(0)t}\mathbf{A}\underline{s}^{(1)} = 0$. Un rapide calcul donne $\mu^{(0)} = \| \mathbf{A}\underline{x}^{(1)} \|^2 \| \mathbf{A}\underline{x}^{(0)} \|^{-2}$.

▷ Méthodes de Newton et quasi-Newton

Nous avons déjà décrit à la section 11.3.3 ("moindres carrés non linéaires") la **méthode de Newton** pour la minimisation d'une fonction $f(\underline{x})$ ($\chi^2(\underline{a})$ dans cet exemple). Après avoir linéarisé au voisinage d'un point $\underline{x}^{(k)}$ choisi proche du minimum l'équation $\underline{\partial} f = 0$, on résoud ce système d'équations linéaires par une méthode itérative. Lorsque la matrice hessienne est définie positive, cette méthode est une généralisation directe de la méthode de Newton-Raphson de recherche de 0 à une dimension (*cf.* section 11.3.2) appliquée à la dérivée $f'(x)$ d'une fonction convexe ($f''(x) > 0$) ; le gradient à n dimensions $\underline{\partial}$ remplace simplement la dérivée première et la matrice hessienne $\mathbf{J}$ la dérivée seconde :

$$f'(x^{(k)}) = -f''(x^{(k)})\,(x^{(k+1)} - x^{(k)}) \quad \longrightarrow \quad \underline{\partial} f(\underline{x}^{(k)}) = -\mathbf{J}^{(k)}\,(\underline{x}^{(k+1)} - \underline{x}^{(k)}) \ .$$

REMARQUE. Lorsque $\mathbf{J}$ est seulement semi-définie positive nous avons introduit la méthode de Levenberg-Marquardt qui consiste à remplacer $\mathbf{J}^{(k)}$ par $\mathbf{J}^{(k)} + \lambda^{(k)}\mathbf{I}$ et à faire tendre $\lambda^{(k)}$ vers 0 lorsqu'on s'approche du minimum. Cette méthode interpole entre celle de la steepest descent (au départ, pour λ grand) et celle de Newton (pour λ petit, près du minimum, lorsque la pente tend vers 0).

La méthode de Newton présente des inconvénients : d'abord il faut déterminer les $\dfrac{n(n-1)}{2}$ dérivées secondes de la matrice hessienne $\mathbf{J}$ et calculer $\mathbf{J}^{-1}$; ensuite le point $\underline{x}^{(k+1)}$ ne correspond pas au minimum de la fonction $f(\underline{x})$ dans la direction $s^{(k)}$ de $\underline{x}^{(k+1)} - \underline{x}^{(k)}$; pire, si $\mathbf{J}$ n'est pas définie positive, ce qui peut arriver loin du minimum, elle peut conduire à s'éloigner de ce minimum (considérer l'exemple $n = 1$ avec $f''(x^{(k)}) < 0$). Dans les **méthodes quasi-Newton** on construit des matrices symétriques définies positives $\mathbf{G}^{(k)}$ jouant le rôle de $[\mathbf{J}^k]^{-1}$ et ne faisant intervenir que $\underline{\partial} f$. Soient $\underline{x}^{(k)}$ et $\mathbf{G}^{(k)}$ le point et la matrice obtenus à l'étape k ; on utilise la méthode de Newton pour déterminer la direction dans laquelle on minimise : $s^{(k)} = -\mathbf{G}^{(k)}\,\underline{\partial} f(\underline{x}^{(k)})$. Soient $\underline{x}^{(k+1)}$ la position du minimum dans cette direction, $\underline{\delta}^{(k)} = \underline{x}^{(k+1)} - \underline{x}^{(k)}$ et $\underline{\gamma}^{(k)} = \underline{\partial} f(\underline{x}^{(k+1)}) - \underline{\partial} f(\underline{x}^{(k)})$; on pose (formule D.F.P. : Davidon, Fletcher, Powell) :

$$\mathbf{G}^{(k+1)} = \mathbf{G}^{(k)} + \frac{1}{\underline{\delta}^{(k)\,t}\underline{\gamma}^{(k)}} \underline{\delta}^{(k)}\,\underline{\delta}^{(k)\,t} - \frac{1}{\underline{\gamma}^{(k)\,t}\mathbf{G}^{(k)}\underline{\gamma}^{(k)}} \mathbf{G}^{(k)}\underline{\gamma}^{(k)}\,\underline{\gamma}^{(k)\,t}\mathbf{G}^{(k)} \ .$$

On itère le processus en partant de $\underline{x}^{(0)}$ et en prenant par exemple $\mathbf{G}^{(0)} = \mathbf{I}$.

REMARQUES. La formule D.F.P. conserve les caractères symétrique et définie positive de $\mathbf{G}^{(0)}$ et assure que $\underline{x}^{(k+1)} - \underline{x}^{(k)} = \mathbf{G}^{(k+1)}\big(\underline{\partial} f(\underline{x}^{(k+1)}) - \underline{\partial} f(\underline{x}^{(k)})\big)$ (condition quasi Newton). Une formule plus élaborée satisfaisant aussi ces conditions est la formule BFGS (de Broyden, Fletcher, Goldfarb et Shanno).

11.5.2 Méthodes du simplex et du recuit simulé

▷ Méthode du simplex

C'est une méthode purement géométrique qui fonctionne en comparant les valeurs de la fonction en différents points qui constituent les vertex d'un simplex. Un **simplex** est une figure géométrique qui, à n dimensions, est constitué de $n+1$ vertex interconnectés par des segments de droites : à deux dimensions ce sont des triangles, à trois dimensions des tétraèdres *etc.* Dans cette méthode la fonction $f(\underline{x})$ étant évaluée aux $n+1$ vertex d'un simplex de l'espace $\{\underline{x}\}$ (les sommets du triangle ABC sur la figure 23 où $n = 2$

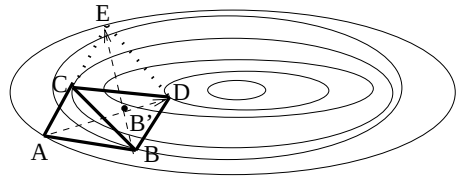


FIGURE 23

et $\underline{x} \equiv (x_1, x_2)$), le vertex qui correspond à la plus grande valeur de f (sommet A) est remplacé par son symétrique (D) par rapport au barycentre des n autres vertex (milieu de BC). On obtient ainsi un nouveau simplex (BCD) sur lequel on recommence l'opération (BCD→CDE) et ainsi de suite. Après p itérations (2 sur la figure), il peut arriver que le nouveau vertex (E) corresponde, comme son prédécesseur (B), à la plus grande valeur de f ; dans ce cas l'itération suivante ramène au simplex de l'étape p (BCD) et la minimisation ne progresse plus. Pour éviter ce phénomène, le simplex est alors contracté dans la ou les directions qui conviennent (par exemple B→B') ; son volume tend donc vers 0 lorsque qu'on se rapproche du minimum.

▷ Méthode du recuit simulé

Cette méthode s'applique à la minimisation d'une quantité $f(\underline{x})$ dépendant de nombreuses variables discrètes ou continues et présentant un très grand nombre de minima secondaires dans lesquels d'autres méthodes risquent de rester "piégées" avant d'atteindre le minimum global. L'idée générale est de générer une marche au hasard (**processus de Markov**) dans l'espace $\{\underline{x}\}$ qui conduit à une **loi de probabilité de Boltzmann** $\exp -\frac{f(\underline{x})}{T}$, puis de faire tendre la "température" T lentement vers 0 pour selectionner la configuration $\underline{x}$ qui minimise $f(\underline{x})$ (figure 24). Si on se trouve dans un minimum secondaire, l'effet de l'"agitation thermique" ($T \neq 0$) est d'en sortir avec une probabilité finie pour explorer l'espace de configuration au delà de ce minima (effet analogue au "recuit" d'un solide qui, préalablement "trempé" dans un état amorphe ou polycristallin correspondant à un minimum local d'énergie, est refondu puis refroidi lentement pour atteindre l'état (mono)cristallin d'énergie minimum).

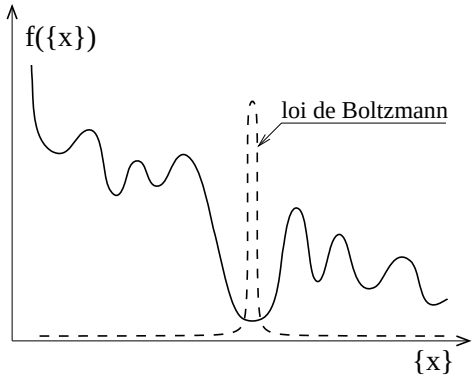


FIGURE 24

EXEMPLES. L'énergie $E(\{S\}) = -\sum_{(ij)} J_{ij} S_i S_j$ de N spins $S_i = \pm 1$ est une fonction qui, pour N grand ($\{S\}$ prend 2^N valeurs) et pour des couplages entre paires de voisins J_{ij} de signes variés, présente un grand nombre de minima. Pour obtenir le minimum global de cette fonction l'**algorithme de Metropolis** consiste, à partir d'une configuration donnée, à tirer un spin i au hasard et à le retourner (changer S_i en $-S_i$) avec une probabilité égale à 1 si le retournement entraîne une diminution de E ($\Delta E < 0$) et égale à $\exp -\dfrac{\Delta E}{T}$ (au lieu de 0) si $\Delta E > 0$. Le lecteur vérifiera que les probabilités de transition ainsi définies entre deux états α et β (qui diffèrent par le retournement d'un spin) vérifient la condition de bilan détaillé $P_{\beta\alpha} \exp -\dfrac{E_\alpha}{T} = P_{\alpha\beta} \exp -\dfrac{E_\beta}{T}$ (*cf.* section 10.3.2) ; ceci garantit que l'algorithme génère une distribution de Boltzmann. En pratique l'algorithme de Metropolis est réalisé en tirant un nombre x au hasard dans $[0, 1]$ (avec une loi uniforme) et en acceptant le retournement si $x < \exp -\frac{\Delta E}{T}$ (quel que soit le signe de ΔE), et en le refusant dans le cas contraire (méthode de rejet). D'autres exemples d'application de la méthode du recuit simulé concerne des problèmes d'optimisation tel que celui du voyageur de commerce, qui consiste à trouver le chemin le plus court passant une seule fois par N villes, ou tel que l'assemblage optimum de circuits élémentaires sur une puce en microélectronique.

Index

aberrations géométriques, 238
accélération, 31, 62, 77
action, 17, **283**, 287, 291, 297, 299
adiabatique (approximation), 195, 298
adimensionnée (équation), 20, 172, 182
Aharonov-Bohm (effet), 293
aire, 26, 61, 69, 93, 208, 221
  loi des, 32
aléatoire, 301
amplitude
  complexe, 39, 42
  de probabilité, 118
analyse
  dimensionnelle, 18
  de Dirac, 150
  de Fourier, 151
  vectorielle, 224
angle, 5, 16
  de diffusion, 37, 116
  solide, 72
anisotrope
  diélectrique, 48, 262
  milieu, 265
antiparticule, 130
antisymétrie, 83, 85
application linéaire, 89, 95
approximation
  adiabatique, 195, 298
  A.R.Q.S., 246, 247, 249, 250
  de Born, 278
  du champ moyen, 143, 313
  $\mathrm{d}_t \simeq \partial_t$, 212
  galiléenne, 55, 113, 291
  de Gauss, 45, **108**, 163
  de la méthode du col, 140, 219
  de l'optique géométrique, 141, **234**
  de Padé, 336
  de la phase stationnaire, 140, 267
  WKB, 197

Babinet (théorème de), 163
barrière de potentiel, 191, 244
base, 88
  changement de, 88, 98
  orthonormée, 90
Bayes (théorème de), 303
Beer (loi de), 7
Bernouilli
  loi de, 304, 307
  théorème de, 71, **213**, 230
Berry (phase de), 299
Bessel (fonctions de), 267
Bethe-Harris (expérience de), 119
Bézier (courbes, surfaces), 341, 342
bifurcation, 180, **197**, 202, 237, 244
  de Hopf, 202
bilan
  détaillé, 320, 321
  d'énergie, 230, 249, 258, 272
  de grandeurs, 227
  de quantité de mouvement, 230, 249
binomiale (loi de probibalité), 310
Biot et Savart (formule de), 211, **246**
Bloch (théorème de), 194
Boltzmann, 8
  E.D.P. de, 272
  loi de, 219, 242, **312**, 319, 352
Born, 87
  approximation de, 278
Bose-Einstein (statistique), 314
boson, 18, 101, 314
boule de billard (mouvement d'une), 82
Bragg (condition de), 85
brownien (mouvement), 21, 160, **318**, 322
bruit
  blanc, 159, **318**, 321
  de grenaille, 312
Buffon (aiguille de), 328
Burgers (E.D.P. de), 274

Cantor, 1
  ensemble de, 21
Cauchy, 3
  loi de, 304
  suite de, 3, 90
Cauchy Riemann (formules de), 279–281

causalité, 112, 155, 270
caustique, 235, 236
Cayley, 87
Cayley-Hamilton (théorème de), 98
cercle, 29, 30
  de Mohr, 107
  osculateur, 62, 137
chaleur, voir thermique
champ
  moyen, 143, 313
  quantique, 129, 130
  vectoriel, 60
changement
  de base, 88, 98
  de jauge, 249, 292
  de référentiel, 5, 53, 77
  d'unité, 2, 15
  de variable, 139, 214, 217, 285
chemin optique, 67, 111, 236, 289
χ^2, 315, 340
Christoffel (symbole de), 296
circuit électrique, 13, 40, 150, 151
circulation, 210, 220, 224, 282
cisaillement, 105
cohérence
  largeur de, 67
  temps de, 159
collisions, 115
commensurables
  fréquences, 158
  grandeurs, 2
commutateur, 95, 97, 120, 124, 127, 287
complexe
  amplitude, 39, 42
  nombre, 22
  notation, 39, 46
  plan, 26
compressibilité, 19, 107, 239, 261
Compton
  effet, 116
  longueur d'onde, 18
concave (fonction), 135, 206
concavité (de l'entropie), 142, 240
conditionnement, 302
conditions initiales (C.I.), 166, 168, 174, 176, 331
conditions aux limites (C.L.), 187, 190, 254, 266
  de Dirichlet, 275, 276, 348
  de Neuman, 275, 276, 348
cône, 72
  de lumière, 112
configuration (espace de), 284
coniques, 32, 37, 99
conjugaison
  complexe, 23, 130
  optique, 108, 189
  de spineurs, 103, 131
contrainte, 9, 242, 285
  extremum avec, 206
  matrice des, 106, 225
contraste, 43, 49
  de phase, 162
convention
  de signe, 11, 14
  de sommation d'Einstein, 94, 224, 294
convolution, 41, **146**, 150, 307
coordonnées
  barycentriques, 27, 338, 339, 349
  curvilignes, 223
  cycliques, 285
  d'espace-temps, 102, 294
  polaires, 31, 184
  sphériques, 71, 224, 278
Copernic (référentiel de), 56, 78
corde de piano, 259
corde vibrante, 256, 269, 271
Coriolis
  accélération de, 77
  force de, 184, 187, 230
corps noir, 9, 19, **239**
corrélation
  fonction de, **40**, 42, 44, 147, 156, 158, 161, 165, **321**
  en probabilité, 308
  temps de, 159, 321
courant de probabilité, 125, 192, 320
courbure
  énergie de, 259
  de l'espace-temps, 294, 296
  de Gauss, 74
  rayon de, 62, 92, 137
  d'une surface, 93

cristal, 84, 256
  uniaxe, 263
crochets de Poisson, 286, 287
Curie
  principe de, 83
  température de, 144
cycle limite, 197, 201, 202

d'Alembert, 253
d'Alembertien, 117
déplacement, 55, 57
de Broglie (relation de), 17, 115, 118, 284
décibel, 6
déformations, 105, 106, 213
dégénérescence (facteur de), 313
degrés de liberté, 88, 124, 168
densité de probabilité, 125, 303
  conditionnelle, 305
  conjointe, 305, 313
  marginale, 305
dérivée, 3, 24, 135, 138
  dans une direction, 209
  au sens des distributions, 149
  logarithmique, 136, 137
  particulaire, 212, 213, 216
  partielle, 204, 212
  totale, 284, 292
Descartes (lois de), 67, 84, 188, 236, 264, 284
détection synchrone, 155, 158
déterminant, 97
développement
  limité, 138, 144
  multipolaire, 91, 104, 251
  de Taylor, 7, 205, 307, 335
diamagnétisme, 293
diélectrique
  anisotrope, 262
  milieu, 197, 247
  polarisation, 155, 198, **247**
différentielle, 66, 205, 215, 219
  de chemin optique, 67
  forme, 219
diffraction
  de Fraunhofer, 46, 162
  de Fresnel, 45, 162
diffusion, 165
  angle de, 37, 116
  E.D.P. de, 262, 268, 271, 317
  état de, 275, 278
  processus de, 317
  de Rutherford, 37, 185
dilatation, 105, 145, 164
dimension, 15
  fractale, 21
dioptre, 30, 67, 73, 109
dipole, 19, 60, 121, 208, 211, 225, 232, 248, 251
Dirac
  analyse de, 150
  E.D.P. de, 131
  impulsion de, 147
Dirichlet (C.L. de), 275, 276, 348
dispersion (relation de), 132, 254, 256, 262, 266
divergence, 210, 211, 223, 272, 286
domaine d'intégration, 216, 218
Doppler (effet), 66, 114
droite, 27
dualité temps fréquence, 154

écart type, 308
échantillonnage, 145, 156, 165
E.D.
  d'amplitude, 195, 199, 200, 202
  d'Euler, 82
  d'Euler-Lagrange, 283, 286, 287, 298
  de Hamilton, 286
  de Schrödinger, 190, 194
E.D.L. du n$^{\text{ème}}$ ordre, 176
E.D.L.S., 25
  du premier ordre, 174
  du second ordre, 25, 174
  avec second membre, 176
  avec second membre sinusoïdal, 40
E.D.P.
  de Boltzmann, 272
  de Burgers, 274
  de diffusion, 262, 268, 271, 317, 333
  de Dirac, 131
  eikonale, 235, 289
  d'Euler
    (hydrodynamique), 230
    (relation), 207, 233, 240

de Fokker-Planck, 272, 318
de Hamilton-Jacobi, 289
de Helmholtz, 275
de Klein-Gordon, 130
de Korteweg-De Vries, 274
de Laplace, 209, 274, 279
de Liouville, 272
de Naviers-Stokes, 231
non linéaire, 273
des ondes élastiques, 231, 262
d'onde simple, 207, 228, 272, 331
de Poisson, 275, 278, 299, 349
de propagation, 215, 262, 269, 332
de Schrödinger, 125, 262, 268, 271, 278, 334
de Sine Gordon, 273
de Weyl, 131
de Weyl-Majorana, 132
Einstein, 51
coefficients d', 321
équations d', 299
relation d' (diffusion), 160, 319
relativité, 111
élasticité, 107, 229, 282
électricité, 12
électromagnétisme, 85, 226, 233, **245**
invariance de l', 63, 117, 297
électrostatique, 31, 151, 209, 231, 246, 274
éléments finis (méthode des), 347
ellipse, 33, 38, 47, 99, 281
émission
induite, 129, 321
spontanée, 129, 321
énergie, 17, 32, 229, 291
bilan d', 230, 249, 258
cinétique, 79, 81, 115, 227, 232
électromagnétique, 130, 233, 250
d'un état stationnaire, 119
flux d', 71, 192, 229, 248, 259
interne, 141, 227, 230, 239, 272, 313
L.C. de l', 123, 258, 266, **289**
de liaison, 19, 115, 291
libre, 143, **240**
potentielle, 32, 125, 170, 181, 231, 291
vitesse de l', 266
ensemble
complet, 3
dense, 3
enthalpie libre, 240
entiers algébriques, 2
entraînement
d'une grandeur, 226, 272
vitesse et accélération d', 77
entropie, **8**, 140, 141, **239**, 272, 313
enveloppe concave, 143, 242
équation
de la chaleur, 272
des cordes vibrantes, 257
d'état, 144, 239
de Langevin, 160, 318
d'Einstein, 299
linéaire, 89, 343
maîtresse, 318, 320
de Maxwell, 117, 245, 248, 262, 297
de Schrödinger, 119, 298
ergodicité, 322
erreur, 309, 315
d'arrondi, 325
de troncature, 325
espace
de configuration, 284
de Hilbert, 90
de phase, 73, 125, **167**, 170, 175, 286, 289
vectoriel (E.V.), 87, 169
dual, 91
des états quantiques, 118
de fonctions, 90
espace-temps, 53, 102, 112, 294
estimateur, 315
état
cohérent, 125, 128
de diffusion, 275
excité, 121
fondamental, 121
gaussien, 125
lié, 190, 194, 275
de polarisation, 46, 96, 118
quantique (E.V.), 118
stationnaire, 120, 188, 298
étendue optique, 73, 109, 190, 286
Euclide, 1, 51
Euler
E.D. d', 82
E.D.P. d' (hydrodynamique), 230

E.D.P. d' (relation), 207, 233, 240
schéma numérique d', 329, 331
Euler-Lagrange (E.D. d'), 283, 286, 287, 298
exponentielle
de $\frac{\mathrm{d}}{\mathrm{d}x}$, 7
fonction, 6, 25
imaginaire, 23
loi de probabilité, 123, 304, 306, 312
de matrice, 96
extremum, 135, 206, 236

Fermat (principe de), 30, 36, **67**, 111, 236, 283
Fermi (règle d'or de), 123
Fermi-Dirac (distribution de), 314
fermion, 18, 101, 131, 314
ferromagnétisme, 143, 315
filtrage, 151, 159
optique, 161
Floquet (théorème de), 193
fluide, 108
fluide parfait, 107, 229, 230
flux, 210, 224
d'un champ de vecteurs, 71
de charges, 70
d'énergie, 71, 192, 229, 248, 259
de masse, 70, 229
de particules, 70
de quantité de mouvement, 70, 229
Fokker-Planck (équation de), 272, 318
fonction
aléatoire, 301, 317
de Bessel, 267
caractéristique, 307, 310, 311
cha ou peigne de Dirac, 148, 152
concave, 135, 206
de corrélation, **40**, 42, 44, 147, 156, 158, 161, 165, **321**
delta, 147
énergie potentielle, 231
exponentielle, 6, 25
gaussienne, **146**, 153, 268
de Green, 268, 270, 271
harmonique, 279
de Heaviside, 148
homogène, 207, 233, 240
hyperbolique, 8
inverse, 137
logarithme, 5
lorentzienne, 146
d'onde, 124
de partition, 313
périodique, 145, 152, 153
porte, 146, 153, 269, 336
propre, 100, 128
sinus cardinal, 41
sinusoïdale, 23
de transfert, 41, 151, 153, 162, 165
de v.a., 304
de z, 24, 279
force, 62, 116, 149
centrale, 32, 36, 38, 80
de Coriolis, 184, 187, 230
de Laplace, 230, **232**, 250
de Lorentz, 117, 292
de pression, 107, 225
forme
différentielle, 219, 239
p-linéaire, 91
quadratique, 91, 111, 178
formule
de Biot et Savart, 211, **246**
de Cauchy Riemann, 279–281
de Green, 277, 349
d'Ostrogradski, 223
de Parseval, 163
de Poisson, 164
de Stirling, 140, 143, 243
de Stokes, 220, 221, 224, 281, 296
Foucault (pendule de), 75, 186
Fourier
analyse de, 151
série de, 152
transformation de, 151
transformée de, 41, 46, 153, 163
foyer (d'une conique), 34
Fraunhofer (diffraction de), 46, 162
fréquence spatiale, 64, 162
Fresnel (diffraction de), 45, 162

Galerkin (méthode de), 347
Galilée, 51
transformation de, 55, 216

Gauss, 22, 51
  approximation de, 45, **108**, 163
  courbure de, 74
  intégration numérique de, 336
  méthode du pivot, 343
  théorème de, 73
gaussien (processus), 321
gaussienne
  fonction, **146**, 153, 268
  loi, 304, 308
gaz parfait, 9, 71, 239, 299
générateur, 8, 96, 97
géodésique, 74, 294
géométrie
  de l'espace, 55
  de l'espace-temps, 294
  de Lobatchevski, 113
  du plan, 26
  sphérique, 16, **74**
  symplectique, 289
Gibbs (phénomène de), 154
Goos-Hänchen (effet), 266
gradient, 115, **208**, 215, 223, 350
grandeur
  additive, 1
  algébrique, 11
  extensive, 9, 19, 240, 309
  intensive, 240
  massique, 141, 227
  molaire, 240
  quadrupolaire, 104, 210, 251
  scalaire, 59, 104
  vectorielle, 59, 104
grandissement (optique), 110
graphe, 134
gravitation, 18, 104, 274, 294
Green
  fonction de, 268, 270, 271
  formule de, 277, 349
grossissement (optique), 110
groupe
  de Lie, 97
  de Lorentz, 102, 103
  des rotations, 103
  de symétrie, 52, 53, 84, 102
gyrotrope (milieu), 48

Hall (effet), 89
Hamilton, 101
  E.D. d', 286
  formule de, 102
  principe de, 286
Hamilton-Jacobi (E.D.P. de), 289
hamiltonien, 119, 286, 287, 292
harmonique, 153, 198
  fonction, 279
harmoniques sphériques, 91, 104, 278
Heisenberg, 87
  inégalité de, 90, 124
hélicité, 118, 131
Helmholtz
  E.D.P. de, 275
  théorème d', 214
Hermite (polynômes de), 128, 334
hermitien (produit), 90
Hilbert, 87
  action de, 299
  espace de, 90
homogénéité, 51, 261, 289
homothétie, 20, 28, 80, 105
Hopf (bifurcation de), 202
Householder (transformation de), 346
Huygens (construction d'), 235
Huygens-Fresnel (principe de), 45, 141, 150, 163
hydrodynamique, 20, 93, 212, 226, 229, 261, 274, 279
hyperbole, 33, 281
hyperboloïde, 36

image
  d'une application linéaire, 89
  méthode des, 276
impédance, 30, 41, **260**, 261
  adaptation d', 191, 260
impulsion de Dirac, 147
incompressible, 213, 279
indépendance
  linéaire, 88, 94
  en probabilités, 303, 308, 310
induction, 12, 226
inégalité
  de Heisenberg, 90, 124
  de Schwarz, 50, 90

inertie
  matrice d', 81
  référentiel d', 51
inertiel (mouvement), 113, 290
inertielle (masse), 290
instabilité, 82, 111, **180**, 195, 197, 200
intégrale première, 32, 38, 168, 184, 285
interactions
  électromagnétiques, 291
  faibles, 63
interférences, 36, **42**, 49, 59, 64, 67, 68, 159, 293
interpolation
  de fonctions, 337
  linéaire, 337
invariance, 15, **52**, 83, 107, 115, 130
  de l'électromagnétisme, 63, 117, 297
inverse (fonction), 137
inversion (géométrie), 28
irréversible, 9, 123, 320
irrotationnel (écoulement), 213, 279
isocline, 171, 175, 200
isotropie, 51, 289

Jacobi (identité de), 95
jacobien, 217
jauge
  changement de, 249, 292
  de Lorentz, 249
  symétrie de, 292

Kelvin (théorème de), 214
Kepler (mouvement de), 36, 184, 295
Klein, 51, 87
Klein-Gordon (E.D.P. de), 130
Koenig (théorème de), 78
Korteweg-De Vries (E.D.P. de), 274
Kramers-Kronig (relations de), 155
Kronecker (symbole de), 90

Lagrange
  multiplicateurs de, 207, 241, 242
  principe de, 284
  théorème de, 214
Lagrange-Helmholtz (relation de), 110
lagrangien, 284
Langevin (équation de), 160, 318
Laplace, 253
  E.D.P. de, 209, 274, 279
  force de, 230, **232**, 250
  transformée de, 123, 157
laplacien, 209, 278, 326
L.C., 85, **228**
  de la charge, 245
  de l'énergie, 123, 258, 266, **289**
  de la masse, 227
  microscopiques, 229
  du moment cinétique, 80, 289
  de la quantité de mouvement, 289
Lebesgue, 133, 160
  mesure de, 3
Le Chatelier (loi de), 243
Legendre (polynômes de), 91, 278, 334
Lenz (loi de), 226
Levenberg-Marquardt (méthode de), 341, 351
Lie, 87
  groupe de, 97
Lienart-Wiechert (potentiel de), 218
lieux géométriques, 30, 35
ligne
  de champ, 168
  électrique, 256
limite
  centrale (théorème de la), 309, 311
  classique, 128
  continue, 257, 317, 318, 320
linéaire
  application, 89, 95
  équation, 89
linéarité, **39**, 41, 147, 150, **169**, 187, 261
Liouville
  E.D.P. de, 272
  théorème de, 286
lissage, 146, 229, 252
Lobatchevski, 51, **113**
localisation (des interférences), 68
logique, 300
loi
  de Descartes, 84, 236
  de Laplace, 93
  de Le Chatelier, 243
  de Lenz, 226
  de Newton, 168, 284
  de Planck, 314, 321

de Stefan, 19, 240
de Van't Hoff, 243
loi de probabilité, 303
de Bernouilli, 304, 307
binomiale, 310
de Boltzmann, 139, 219, 242, **312**, 319, 352
de χ^2, 315, 340
exponentielle, 123, 304, 306, 312, 327
gaussienne, 139, 304, 308, **308**, 327
lorentzienne, 304, 308
de Poisson, 129, **311**
uniforme, 304, 306, 307, 327
Lorentz
condition de, 117
force de, 117, 292
groupe de, 102, 103, 295
jauge de, 249
transformations de, 54, 112, 216
lorentzienne
fonction, 146
loi, 304, 308
Lotka-Volterra (oscillateur de), 171

magnétostatique, 31, 89, 246, 251, 274
Magnus (effet), 280
Malus (théorème de), 68
marche au hasard, 309, 317
marée (champ de), 295
Markov (processus de), 319
masse
défaut de, 115
et fréquence, 132
gravitationnelle, 294
inertielle, 132, 290
nulle, 115, 131
réduite, 79
matrice, 94
adjointe, 94
colonne, 94
des contraintes, 106, 225
des déformations, 105
densité, 50, 119
diagonale, 99, 347
exponentielle de, 96
hermitienne, 50, 95, 96, 99, 102
hessienne, 290, 340
d'inertie, 81
inverse, 97
de Jones, 49
de Jordan, 99
orthogonale, 95, 96, 100, 215, 347
de Pauli, 101, 102
de pression, 106, 249
projecteur, 95, 101
S, 192
spectre d'une, 98
symétrique, 95, 99, 105, 178, 262, 310
symplectique, 289
de transfert, **108**, 110, **188**, 203
de transition (probabilité), 319
transposée, 94
triangulaire, 344, 347
tridiagonale, 334, 344, 346, 348
unitaire, 95, 96, 99, 102, 192
Maupertuis (principe de), 284
maximum de vraisemblance, 315
Maxwell, 253
construction de (palier gaz-liquide), 242
distribution de, 137, 272, 306
équations de, 117, 245, 248, 262, 297
mécanique du solide, 76
Meissner (effet), 293
métastabilité, 143, 244
méthode
du col, 140
des isoclines, 200
de la phase stationnaire, 140
méthode (numérique)
de Bézier, 341
de collocation, 350
des éléments finis, 347
de Galerkin, 347
du gradient conjugué, 350
de Householder, 346
d'intégration de Gauss, 336
de Lax, 332
de Levenberg-Marquardt, 341, 351
LU, 344
des moindres carrés, 339
de Monte-Carlo, 328
de Neville, 338
de Newton, 340, 351
de Newton-Raphson, 339

du pivot de Gauss, 343
QR, 347
quasi Newton, 351
du recuit simulé, 352
de rejet, 328
de la sécante, 339
du simplex, 352
spectrale, 347
de la steepest descent, 350
SVD, 345
de Tikhonov, 346
de tir, 331
du trapèze, 326
métrique, 91
de l'espace-temps, 111, 294
de la sphère, 72
structure, 90
métrologie, 17
Metropolis (algorithme de), 353
microréversibilité, 123, 320
Miller (indices de), 65
minima (recherche de), 350
miroir, 36, 59, 67, 109, 111
Möbius (bande de), 222
modes propres, 100, 177, 179, 255
modulation, 195
d'amplitude, 42, 155, 160
de fréquence, 42, 160
module de rigidité, 107
moindres carrés (méthode des), 339
moment
cinétique, 69, 79, 81, 118, 185
cinétique (L.C. du), 80, 289
conjugué, 286, 292
de la loi de Gauss, 139
magnétique, 69, 185, 209, 248, 251
d'une v.a., 307
mouvement
d'une boule de billard, 82
brownien, 21, 160, **318**, 322
harmonique, 38, 184
inertiel, 113, 290
de Kepler, 36, 184
d'une particule chargée, 116, 186, 196, 292
de précession, 185
d'un solide, 58, 80
de translation, 56
uniformément accéléré, 182
moyenne, 307
quantique, 119
temporelle, **39**, 119, **158**, 161, 322
moyennisation, 41, 146
multiplicateurs de Lagrange, 207, 241, 242
multipolaire (développement), 91

Naviers-Stokes (E.D.P. de), 231
Neuman (C.L. de), 275, 276, 348
neutrino, 131
Neville (méthode d'interpolation), 338
Newton
loi de, 168, 284
méthode de minimisation de, 340, 351
relations de (optique), 110
Newton-Raphson (méthode de), 339
nombres
aléatoires, 327
algébriques, 2
complexes, 22
rationnels, 2
réels, 3
notation complexe, 39, 46
noyau (d'une application), 89

Ohm (loi d'), 13
onde
de choc, 207, 228, 229, 274
élastique, 231, 262
évanescente, 191, 192, 264
guidée, 266
et mécanique, 236, 284, 286
plane, 64, 161, 235, 261
propagative, 191, 254
simple (E.D.P. d'), 207, 228, 272
sismique, 262
solitaire, 273
sonore, 19, 86, 188, 212, 229, **261**
sphérique, 67, 150, 267, 270
stationnaire, 190, 254, **271**
de surface, 261, 274
surface d', 141, 234
opérateur
d'annihilation, 128, 130
de création, 128, 130
hamiltonien, 119

laplacien, 209, 278
nabla, 209, 211
de position, 124
quantité de mouvement, 124
de symétrie, 118, 120, 124
optique
de Fourier, 161
géométrique (approximation de l'), 141, **234**
incohérente, 151, 165
matricielle, 108, 111, 127
non linéaire, 198
orbitale atomique, 104
oscillateur
amorti, 149, 174
couplés, 92, 100, 178, 200
entraîné avec frottement solide, 172
harmonique quantique, 126, 127
de Lotka-Volterra, 171
non linéaire, 197
paramétrique, 173, 187, 193, 195, 285
de Van der Pol, 20, 200, 202
Ostrogradski (formule de), 223

paquet d'ondes, 124, 127, **141**, **264**, 274, 286
parabole, 33
paramagnétisme, 143, 293, 315
paramétrage additif, 4
paramètre d'ordre, 244, 313
paramétrique (oscillateur), 173, 174, 187, 193, 195, 198, 285
paramétrique (oscillateur amorti), 195
parité, 55, 145
Parseval (formule de), 156, 163
partition
de l'ensemble des résultats, 302
fonction de, 313
Pauli
matrices de, 101, 102
principe de, 131
pendule de Foucault, 75, 186
percussion, 82, 124, 149, 269
périodique (fonction), 145, 152, 153
périodisation, 145, 156, 164
permutation, 94, 101
Perot Fabry, 44
phase
de Berry, 299
espace de, 125, 167, 286, 289
vitesse de, 64
phonon, 19, 256
photon, 73, **96**, 116, **118**, 119, 123, **129**, **314**
plan, 64
complexe, 26
réticulaire, 65, 85
tangent, 206, 208
Planck
échelle de, 18
facteur de, 74
loi de, 137, **314**, 321
Planck-Einstein (relation de), 17, 115, 118
point fixe, 171
points de vue actif et passif, 53
Poisson, 253
coefficient de, 108
crochets de, 286, 287
E.D.P. de, 275, 278, 299, 349
formule de, 152, 154, 164
loi de, 129, **311**
polarisation, 38, 100, 187
diélectrique, 155, 198, **247**
états de, 46, 96, 118
polynôme
de Berstein, 341
caractéristique, 98
de Hermite, 128, 334
de Legendre, 91, 278, 334
minimal, 98
de Tchebytchev, 335
porte (fonction), 146
portrait de phase (définition), 170
potentiel
chimique, 240, 242, 313, 314
de Lienart-Wiechert, 218
scalaire, 248, 292
thermodynamique, 243, 313
vecteur, 211, 248, 292
Poynting (vecteur de), 248
précession (mouvement de), 185
pression, 239
cinétique, 71, 196, 299
force de, 107, 225

de radiation, 71, 196
primitive, 136, 138
principe
de Curie, 83
de Fermat, 30, 36, **67**, 111, 236, 283
de Hamilton, 286
de Huygens-Fresnel, 45, 150, 163
de Lagrange, 284
de Maupertuis, 284
de moindre action, 283
de Pauli, 131
probabilité, 302
amplitude de, 118
courant de, 125, 192, 320
densité de, 125, 303
lois de, 303
de transition, 123, 129, 317, 319
problème à deux corps, 79
processus
aléatoire, 301, 317
de diffusion, 317
gaussien, 321
de Markov, 319
poissonnien, 312, 321
stationnaire, 321
produit
hermitien, 90
mixte, 60, 93
scalaire, 26, **60**, **90**, 95
tensoriel, 88, 104
vectoriel, 60
projecteur, 50, **90**, 95, 101
propagation (E.D.P. de), 215, 262, 269
pseudovecteur, 62
puits de potentiel, 120, **190**, 192, 194
Pythagore (théorème de), 16

QR (algorithme), 346
quadrivecteur, 114, 216, 296
énergie-quantité de mouvement, 114
quadrupole, 19, 104, 251
quantification, 190, 194, 266, 278
quantité de mouvement, 79, 124, 227, 291
L.C. de la, 289
quasi-monochromatique (signal), 42, 123, 159
quasi-périodique (signal), 158

racines d'un polynôme, 25
rang (d'une application linéaire), 89
rapidité, 5, 54, 112
Rayleigh-Ritz, 100
rayon
de courbure, 62, 137
optique, 67, 108, 188, 189, **234**, 235, 296
rayonnement, 19, 86, 251
recuit simulé (méthode du), 352
récurrence (relation de), 254
redshift gravitationnel, 295
référentiel
du centre de masse, 78, 115
changement de, 53, 77
de Copernic, 56, 78
géocentrique, 56
d'inertie, 51
local, 78, 183
réflexion
coefficient de, 191, 260, 266
totale, 264, 266
règle d'or de Fermi, 123, 129
relativité d'Einstein, 17, 111
renversement du temps, 55
réponse
impulsionnelle, 149, 150
linéaire (Boltzmann), 314
représentation linéaire (d'un groupe), 52, 102–104
réseau, 43, 65, 84, 163, 164
réciproque, 65
résonance, 40, 179
non linéaire, 199
paramétrique, 174, 193, 195, 198
quantique, 122
résultats (probabilités), 301
réversible, 55, 262, 298
Reynolds, 15
nombre de, 20
Ricci (courbure de), 296
Richardson (extrapolation de), 327
Riemann, 51
courbure de, 296
intégrale de, 3, 216
R.M.N., 177
rotation, 28, 54, **56**, 58, 60, **76**, 96, 213, 231
symétrie de, 85, 288

rotationnel, 210, 211, 223
Runge-Kutta (schéma numérique de), 329
Rutherford (diffusion de), 37, 185

scalaire
  grandeur, 59, 104
  potentiel, 292
  produit, 60, 90, 95
schéma numérique
  de Crank-Nicolson, 329, 333
  d'Euler, 329, 331
  leapfrog, 330, 332
  de Runge-Kutta, 329
Schrödinger, 87, 253
  E.D. de, 188, 190, 194
  E.D.P. de, 125, 262, 268, 271, 278
  équation de, 119, 298
Schwarz
  égalité de, 205, 236, 240
  inégalité de, 50, 90, 147
Schwarzschild
  action de, 297
  métrique de, 295
séparation des variables, 278
série de Fourier, 152
Shannon (condition de), 156, 165
signal
  chaotique, 158, 159
  complexe (analytique), 39
  quasi-monochromatique, 42, 123, 159
  quasi-périodique, 158
  stationnaire, 158
signal lumineux, 156, 159, 312, 321
Sine Gordon (E.D.P. de), 273
Smith (abaque de), 30
solide, 9
  champ des vitesses d'un, 80, 211
  mécanique du, 76
  mouvement d'un, 58, 80
somme directe (d'E.V.), 88, 104
spectre
  de fréquences d'un signal, 153
  d'une matrice, 98
  d'une v.a., 301
sphère, 72
spineur, **103**, 104, 131
splines
  B-splines, 336
  cubic-splines, 337
stabilité
  numérique, 325, 330, 332, 333
  d'un système dynamique, 172, 180
  thermodynamique, 241
stationnarité, **41**, 147, 150, 168, 212, 216, **261**, **289**
statistiques quantiques, 314
Stefan (loi de), 19, 240
Stirling (formule de), 140, 143, 243
Stokes (formule de), 220, 221, 224, 281, 296
supraconducteur, 277, 293
surface, 206, 208, 212
  fermée, 224
  de niveau, 208
  d'onde, 141, 234
  orientable, 221
  ouverte, 224
  des vecteurs d'onde, 263
SVD (singular value decomposition), 345
symétrie, 52, 288
  discrète, 58, 62, 85, 95
  groupe de, 52, 53, 84
  de jauge, 292
  opérateur de, 118, 120, 124
  de rotation, 85, 285, 288
  de translation, 84, 288
symbole de Kronecker, 90
symplectique (géométrie ; matrice), 289
système dynamique, 167, 170, 207

Taylor (développement de), 7, 135, 205, 307, 335
Tchebytchev (polynômes de), 335
température, 17, 239
temps
  de cohérence, 159
  de corrélation, 159, 321
  propre, 113, 290, 294
tension (d'un ressort), 12
théorème
  de Bayes, 303
  de Bernouilli, 71, **213**, 230
  de Bloch, 194
  de Floquet, 193
  d'Helmholtz, 214

de Kelvin, 214
de Lagrange, 214
de la limite centrale, 309, 311
de Liouville, 286
du viriel, 233
de Wiener-Khintchine, 158, 322
thermique (échange), 71, 229, 239, 298
thermodynamique, 27, **141**, **238**
tire-bouchon (règle du), 13, 57, 222
trace, 97
transfert
fonction de, 41, 151, 153, 162, 165
matrice de, **108**, 110, **188**, 203
transformation
canonique, 287, 290
conforme, 29, 282
de Fourier, 151
de Galilée, 55, 216
de Lorentz, 54, 102, 112, 216
transformée
de Fourier, 41, 46, 153, 163
de Fourier discrète, 157, 334
de Laplace, 123, 157
en z, 24, 41
transition
de phase, 142
quantique, 122
translation, 28, 56, 60, 145, 164, 231
d'espace-temps, 54
mouvement de, 56
symétrie de, 84, 288
transmission (coefficient de), 191
transport
parallèle, 74, 296
quantique, 299
travail, 231, 233, 250
en thermodynamique, 142, 243, 298
travaux virtuels (théorème des), 286

v.a. indépendantes, 305, 307, 308, 317
valeur propre, 98–101, 106, 111, 121, 177, 319, 346
Van der Pol (oscillateur), 20, 200, 202
Van der Waals
équation de, 135, 239
fluide de, 241, 244
Van't Hoff (loi de), 243
variance, 307, 309
vecteur
axial, 63
de Jones, 47
de Lenz, 37, 184
d'onde (surface des), 263
polaire, 62
potentiel, 292
de Poynting, 71, 248
propre, 98–101, 121, 177, 319, 345, 346
surface, 69
tournant, 38, 76, 185, 187
unitaire, 31, 62
vectoriel (produit), 60
vectorielle (grandeur), 59
viriel (théorème du), 233
viscosité, 108, 282
vissage, 57
vitesse
aréolaire, 27
c (de la lumière), 54
de l'énergie, 266
de groupe, 125, 132, 141, 236, 265, 286
de phase, 64
volume, 61, 106, 213
forme, 93
à très grand nombre de dimensions, 218

Weyl (E.D.P. de), 131
Wiener-Khintchine (théorème de), 158, 322
WKB (approximation), 197
wronskien, 189, 192

Young (module d'), 108, 258

Zeeman (effet), 186